Universitat Autònoma de Barcelona

Audio subtitling: Voicing strategies and their effect on emotional activation

Gonzalo Iturregui-Gallardo

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: http://es.creativecommons.org/blog/licencias/

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/?lang=en

PhD in Translation and Intercultural Studies

Audio subtitling:

Voicing strategies and their effect on emotional activation

PhD Thesis presented by:
Gonzalo Iturregui-Gallardo

Supervised by:
Dr. Anna Matamala

February 2019

Department of Translation and Interpreting
and East Asian Studies

*Departament de Traducció i d'Interpretació
i d'Estudis de l'Àsia Oriental*

Universitat Autònoma
de Barcelona

Universitat Autònoma de Barcelona

Departament de Traducció i d'Interpretació i d'Estudis de l'Àsia Oriental
Doctorat en Traducció i Estudis Interculturals

Department of Translation and Interpreting and East Asian Studies
PhD in Translation and Intercultural Studies

Audio subtitling:

Voicing strategies and their effect on emotional activation

PhD Thesis presented by:

Gonzalo Iturregui-Gallardo

Supervised by:

Dr. Anna Matamala

February 2019

To Miguel

Acknowledgements

First of all, for the opportunity to conduct this research project I would like to express my gratitude to the Secretaria d'Universitats i Recerca of the Catalan Government (2016FI_B 00012) and the project Nuevos Enfoques sobre Accesibilidad (FFI2015-64038-P MINECO/FEDER, UE) and its team for providing me with enough funding and resources.

I am very grateful to the Associació Discapacitat Visual Catalunya B1, B2, B3 and its members for its collaboration in the experiments.

Enormes gracias a mi familia, que siempre me ha animado a seguir para conseguir todos mis objetivos y ha sabido guiarme a través de todas mis idas y venidas que me han llevado a emprender este proyecto.

Gracias a mi equipo de biblioteca, mis futuros médicos, que me han acompañado durante horas entre libros, notas y muchos subralladores.

It's been a pleasure to share my advances and challenges with my MRA colleagues, especially those who have already obtained their "Dr." before their names. Also, thank you Anja for being such a help in and out the thesis.

No puc deixar de mencionar a la colla que porta acompanyant-me tota la vida, sempre de la mà, i ha sigut partícep de totes les decisions que m'han portat on sóc ara.

Rigranzio ancora Brema e tutto ciò che mi ha datto. I nostri ritrovi sono stati le migliori pause per prendere aria e riprendere questa avventura

I give my thanks to the city of Krakow, that has been an unexpected bonus of this thesis. *Dziękuję* to Michal, Joana and Kinga, who have all become very important parts of the thesis and my experience in such a wonderful city.

I could not go without mentioning Anna Lyons, who has dealt with all my English writing, a sincere thanks.

And finally, my special thanks to my supervisor, Dr. Anna Matamala. Without her knowledge, perseverance and professionality the finalisation of this thesis would have not been possible. Thank you for having given me the opportunity for joining the research group and complete this PhD thesis.

Abstract

Blind and partially sighted audiences have limited access to conventional audiovisual materials. Images must be adapted into audio format in order to provide the user with enough information for the correct understanding and enjoyment of the plot. Audio description (AD) stands as the most popular way to overcome such barriers. However, what happens with subtitled materials? The response has been found in the use of audio subtitling (AST), which has been described as the intersection between subtitles, AD and voice-over. Despite the clear need for them, they have only been used regularly by a few countries on TV since the 90s. Although the main focus of attention has been put on the study of the technology used in the delivery of computer AST less has been said about human voiced AST, the actual reception of AST in fiction films and how AST has to be merged with AD and presented to the audience. This research is carried out within the framework of Audiovisual Translation Studies and Media Accessibility and it borrows knowledge from other fields such as psychology. It aims to shed a light on the impact on the audience of two different AST strategies: dubbing effect and voice-over effect. These two strategies for the delivery of AST will be compared in terms of emotional activation by making use of psychophysiological measures, which are able to measure physiological responses such as electrodermal activity or heart rate, in combination of self-report instruments and the study of preferences.

Keywords: media accessibility, audio subtitling, audio description, emotion, psychophysiology

Resumen

El público ciego o con baja visión tiene un acceso limitado a los materiales audiovisuales convencionales. La imagen debe ser adaptada a un formato de audio capaz de dotar al espectador con la suficiente información para que haya una correcta comprensión y disfrute del contenido. La audiodescripción (AD) es la estrategia más conocida para llevar a cabo este tipo de adaptación. Sin embargo, ¿qué sucede con los contenidos subtitulados? La respuesta se encuentra en el uso del audiosubtitlulado, que se describe como el cruce entre subtítulo, AD y voice-over. Aún siendo muy necesarios, tan solo algunos países lo utiliza con regularidad en TV desde los 90. Aunque el estudio del AST se centra principalmente en las tecnologías para su aplicación informatizada, poco se ha dicho sobre el AST narrado con voces humanas y la forma en que estos se perciben en películas, se combinan con la AD y se presentan al público. Esta tesis se sitúa en el marco de la Traducción Audiovisual y la Accesibilidad a los Medios de Comunicación abarcando, asimismo, disciplinas como la psicología. El objetivo principal es esclarecer el impacto que tienen los dos diferentes modos de narrar el AST: con efecto de doblaje y con efecto de voice-over. Dichos modos se compararan teniendo en cuenta la activación emocional mediante el uso de medidas psicofisiológicas, capaces de medir las respuestas fisiológicas del organismo, como la actividad electrodérmica y la frecuencia cardiaca, el uso de cuestionarios y el estudio de preferencias.

Palabras clave: accesibilidad a los medios audiovisuales, audiosubtitulado, audiodescripción, emoción, psicofisiología

Table of Contents

CHAPTER 1 INTRODUCTION 3

1. OBJECTIVES AND HYPOTHESES 4
2. THEORETICAL FRAMEWORK 6
3. METHODOLOGY 7
4. THESIS STRUCTURE 8

CHAPTER 2 THEORETICAL FRAMEWORK 15

1. AUDIOVISUAL TRANSLATION AND MEDIA ACCESSIBILITY 15
2. THE USER: ACCESSIBILITY FOR VISUAL AND READING DISABILITIES 17
3. DEFINING AST 18
   3.1. AD: Transforming Images into Words 19
   3.2. AST: Converting Written Text into Sound 20
4. CATEGORISING AUDIO SUBTITLES 21
   4.1. Combination: Autonomous and Integrated 22
   4.2. Voices: Synthetic and Human 23
   4.3. Mix: Broadcast-Mix and Receiver-Mix 24
5. RESEARCH ON AST 24
6. AST AND MULTILINGUALISM 28
   6.1. Revealing Multilingualism with AST 29
7. DUBBING AND VOICE-OVER EFFECT 32
8. REGULATING MEDIA ACCESSIBILITY: AST AND AD 36
9. STANDARDS AND GUIDELINES: AST AND AD 42
10. CONCLUSION 45

CHAPTER 3 USER RECEPTION: THEORIES AND METHODS 51

1. USER RECEPTION IN AUDIOVISUAL TRANSLATION AND MEDIA ACCESSIBILITY 51
2. MEDIA AS ENTERTAINMENT 52
   2.1. Media Entertainment: Flow 54
   2.2. Media Entertainment: Presence 55
3. ENTERTAINMENT, FLOW, PRESENCE AND EMOTIONAL ACTIVATION 57
4. EMOTIONS 59
   4.1. Defining Emotion 60
   4.2. Categorisation of Emotions: Discrete and Dimensional Approaches 62
   4.2. Emotions Put to the Test 70
5. MEASURING TOOLS 76
   5.1. Self-Report Instruments 76
   5.2. Measuring Physiological Response 79
6. FROM QUESTIONNAIRES TO PHYSIOLOGY IN AD AND AST RESEARCH 82
7. CONCLUSION 86

CHAPTER 4 METHODOLOGY OF STUDY 1: DESCRIPTIVE ANALYSIS 91

1. MATERIALS: CORPUS SELECTION 91
2. METHODOLOGY: DATA EXTRACTION AND ANALYSIS 98

CHAPTER 5 DISCUSSION OF RESULTS IN STUDY 1 107

1. MATRIX OF ANALYSIS: GENERAL AST CHARACTERISTICS 107
2. DISCUSSION 115
   2.1. Presence of Audio Subtitles 115

2.2. Voices *116*
2.3. Audibility *116*
2.3. Isochrony *117*
2.4. Prosody *121*
2.5. Assignation *122*
2.6. Textual Features *123*
3. CONCLUSIONS 124

CHAPTER 6 STUDY 2: MEASURING USER EXPERIENCE 131

1. PHASE 1: ONLINE VALIDATION 132
1.1. Participants *132*
1.2. Materials *132*
1.3. Duration *137*
1.4. Ethical Permission *137*
1.5. Contacting the Respondents *137*
1.6. The Tasks *138*
1.7. Pilot Validation *138*
1.8. Final Clips *139*
2. PHASE 2: THE EXPERIMENT 140
2.1. Experiment 1 .... *140*
2.2. SAM Questionnaire for Blind and Partially Sighted Participants *149*
2.3. Experiment 2 .... *157*
3. DATA ANALYSIS 163
3.1. Self-Report Measures: T-SAM *163*
3.2. Psychophysiological Measures *163*

CHAPTER 7 DISCUSSION OF RESULTS IN STUDY 2 169

1. EXPERIMENT 1 170
1.1. Results for the Self-Report Questionnaire *170*
1.2. Results for Psychophysiological Measures *176*
1.3. Results for the Preference Questionnaire *181*
2. EXPERIMENT 2 188
2.1. Results for the Self-Report Questionnaire *188*
2.2. Results for the Psychophysiological Measures *191*
3. COMPARISON OF RESULTS IN EXPERIMENTS 1 AND 2 193
3.1. T-SAM Questionnaire *193*
3.2. Psychophysiological Measures *194*
4. CORRELATIONS 196
4.1. Self-Report and Psychophysiological Results *197*
4.2. Psychophysiological Results: EDA and HR *199*

CHAPTER 8 CONCLUSIONS 203

1. VOICE-OVER EFFECT VS. DUBBING EFFECT: EMOTIONAL ACTIVATION AND PREFERENCES 203
2. UNDERSTANDING AUDIO SUBTITLING: DEFINITION AND CURRENT PRACTICES 206
3. DEFINING A THEORETICAL FRAMEWORK: AUDIOVISUAL TRANSLATION, MEDIA ACCESSIBILITY AND USER EXPERIENCE 210
4. COMPARING THE EMOTIONAL ACTIVATION OF DIFFERENT USER PROFILES 211
5. MAIN CONTRIBUTIONS AND LIMITATIONS 213

REFERENCES 221

FILMOGRAPHY 257

ANNEXES I

1. SUMMARY OF REGULATIONS, STANDARDS AND GUIDELINES I
2. DOCUMENTS RELATED XIII
2.1. Phase 1: Online Validation xiii
2.2. Phase 2: T-SAM Validation (Focus Group) xv
2.3. Phase 2: Experiment 1 xviii
2.4. Phase 2: Experiment 2 xxvi

List of Figures

Figure 1. From entertainment to emotion 69
Figure 2. Percentage of audio subtitles for type of isochrony in *Everything Is Illuminated* 117
Figure 3. Percentage of audio subtitles for each type of isochrony in *The Science of Sleep*...118
Figure 4. Percentage of audio subtitles for each type of isochrony in *Syriana* 119
Figure 5. Percentage of audio subtitles for each type of isochrony in *Slumdog Millionaire* 119
Figure 6. Percentage of audio subtitles for each type of isochrony in *Borat*...120
Figure 7. Percentage of audio subtitles for each type of isochrony in *Munich*120
Figure 8. Films organised in terms of acting in the AST Service 126
Figure 9. Valence dimension of the SAM questionnaire................................137
Figure 10. Arousal dimension in the SAM questionnaire...............................137
Figure 11. T-DAC, receiver device of the Captiv L7000 141
Figure 12. GSR sensor of the Captiv L7000...141
Figure 13. HR sensor of the Captiv L7000 142
Figure 14. Tobii eye-tracker receiver and sensor 142
Figure 15. Toshiba Portégé 730-B-119 143
Figure 16. HP Elitedisplay E232 monitor 143
Figure 17. Sony MDR-ZX110NA head-phones 144
Figure 18. Example of the Google Forms sheet 144
Figure 19. Version 1 of the T-SAM 152
Figure 20. Version 2 of the T-SAM 152
Figure 21. Final version of the T-SAM 157
Figure 22. Structure of Study 2..162
Figure 23. Valence ratings for emotion (Experiment 1): median, average (standard deviation)..171
Figure 24. Valence ratings for the AST effect (Experiment 1): median, average (standard deviation)..172
Figure 25. Valence ratings for emotion and AST effect (Experiment 1): media, average (standard deviation)...173
Figure 26. Arousal ratings for the emotion (Experiment 1): median, average (standard deviation)..174
Figure 27. Arousal ratings for the AST effect (Experiment 1): median, average (standard deviation)..175
Figure 28. Arousal ratings for emotion and AST effect (Experiment 1): median, average (standard deviation)...176
Figure 29. EDA stimuli averages minus 3-minute baseline average for emotion (Experiment 1) 177
Figure 30. EDA stimuli averages minus 3-minute baseline average for AST effect (Experiment 1) 178

Figure 31. EDA stimuli averages minus 3-minute baseline average for emotion and AST effect (Experiment 1) 178
Figure 32. HR stimuli averages minus 3-minute baseline average for emotion (Experiment 1) 179
Figure 33. HR stimuli averages minus 3-minute baseline average for AST effect 180
Figure 34. HR stimuli averages minus 3-minute baseline averages for emotion and AST effect (Experiment 1) 181
Figure 35. Replies to “Which one of the two options do you prefer?”............183
Figure 36. Replies to “Which one do you think can transmit better the emotion of the scene?”...184
Figure 37. Replies to “Was there an option in which dialogues where easier to understand?” 185
Figure 38. Replies to “Which kind of voice do you prefer for the AST (when subtitles are read aloud)?”..186
Figure 39. Replies to “Would you prefer having a different voice for each character?” 187
Figure 40. Replies to “Would you like to have a more acted audio description?” 188
Figure 41. Valence ratings for emotion (Experiment 2): median, average(standard deviation)...189
Figure 42. Arousal ratings for emotion (Experiment 2): median, average(standard deviation)...190
Figure 43. EDA stimuli averages minus 3-minute baseline average for emotion (Experiment 2) 191
Figure 44. HR stimuli average minus 3-minute baseline average for emotion (Experiment 2) 192
Figure 45. Valence ratings for emotion (Experiment 1 and 2): median, average(standard deviation)...193
Figure 46. Arousal ratings for emotion (Experiment 1 and 2): median, average(standard deviation)...194
Figure 47. EDA stimuli averages minus 3-minute baseline average for emotion (Experiment 1 and 2) 195
Figure 48. HR stimuli averages minus 3-minute baseline average (Experiments 1 and 2) 196

List of Tables

Table 1. Characteristics of the emotions: dimensional approach and physiological responses 75
Table 2. Self-report instruments on emotional activation through audiovisual stimuli 79
Table 3. General information about *Everything Is Illuminated* 94
Table 4. General information about *The Science of Sleep* 95
Table 5. General information about *Syriana* 96
Table 6. General information about *Borat* 97
Table 7. General information about *Munich* 97
Table 8. General information about *Slumdog Millionaire* 98
Table 9. Summary matrix for the results of Study 1 103
Table 10. Details of *The Science of Sleep* 108
Table 11. Details of *Everything Is Illuminated* 109
Table 12. Details of *Syriana* 110
Table 13. Details of *Slumdog Millionaire* 111
Table 14. Details of *Borat*...112
Table 15. Details of *Munich*...113
Table 16. Summary of the features in the matrix of analysis. 114
Table 17. Summary of the scenes identified for the online validation 136
Table 18. Online validated scenes and their ratings 139
Table 19. Experimental clips 146
Table 20. Randomisation of the stimuli presentation 147

Abbreviation and Acronym Glossary

2D	2 Dimensions
3D	3 Dimensions
AD	Audio Description
AD-ACL	Activation-Deactivation Adjective List
ADLAB	Audio Description: Lifelong Access for the Blind
AENOR	Asociación Española de Normalización y Certificación
AI	Audio Introduction
AMI	Accessible Media Inc.
ANOVA	Analysis of Variance
AST	Audio Subtitling
AVT	Audiovisual Translation
BBC	British Broadcasting Company
bpm	Beats per minute
BVP	Blood Volume Pulse
CAB	Canadian Association of Broadcasters
DR1	Danmarks Radio 1
DR1Syn	Danmarks Radio 1 Syn
Dub	Dubbing
DVD	Digital Versatile Disc
EDA	Electrodermal Activity
EEG	Electroencephalography
ELJABR	Elimination of the Languages Barriers Faced by the Handicapped Watchers of the Czech Television
EU	European Union
fMRI	Functional Magnetic Resonance Imaging
FNB	Federation of Dutch Libraries for the Blind
FSB	Netherlands Federation of the Blind and Partially Sighted
H1	Hypothesis 1
H2	Hypothesis 2
H3	Hypothesis 3
H4	Hypothesis 4

HR	Heart Rate
HRV	Heart Rate Variability
IEC	International Electronic Commission
ISO	International Organisation for Standardisation
ITC	Independent Television Commission
MA	Media Accessibility
MAA	Media Access Australia
MAACL	Multiple Affect Adjective Checklist
MAACL-R	Multiple Affect Adjective Checklist Revised
MACL	Mood Adjective Checklist
Min	Minutes
NEA	Nuevos Enfoques sobre Accesibilidad
NOS	Nederlandse Omroep Stichting
OCR	Optical Character Recognition
Ofcom	Office of Communications
ONLUS	Organizzazione Non Lucrativa di Utilità Sociale
PACMA	Perception, Emotion, Cognition and Motor Activation
PANAS	Positive and Negative Affect
PANAS-X	Positive and Negative Affect Schedule
PNG	Portable Network Graphics
POMS	Profile of Moods Stats
RP	Received Pronunciation
RTP1	Rádio e Televisão de Portugal 1
RTP2	Rádio e Televisão de Portugal 2
S4C	Sianel Pedwar Cymru
SAM	Self-Assessment Manikin
SDH	Subtitles for the Deaf and Hard-of-Hearing
Sec	Seconds
SIC	Sociedade Independente de Comunicação
STAI	State-Trait Anxiety Inventory
T-DAC	TDM-Digital Access Concentrator
T-SAM	Tactile Self-Assessment Manikin
TAS-20	Toronto Alexithymia Scale

TEA	Tech Ergo Appliquées
TTS	Text-to-Speech
TV	Television
TVI	Televisão Independente
TVP1	Telewizja Polska 1
UAB	Universitat Autònoma de Barcelona
UK	United Kingdom
UKAAF	UK Association for Accessible Formats
UNE	Asociación Española de Normalización
UX	User Experience
VE	Virtual Environment
VGA	Video Graphics Array
VO	Voice-over
VR	Virtual Reality
ZKPQ	Zuckerman-Kuhlman Personality Questionnaire
μS	Micro Siemens

Chapter 1

Introduction

Chapter 1
Introduction

An audio subtitle can be defined as an aurally rendered version of a written subtitle (Reviers & Remael, 2015). Although primarily targeted at blind and partially sighted users as a part of audio description (AD) or independently, audio subtitling (AST) can also enhance the audiovisual experience to people with diverse needs, such as persons with reading disabilities or second language learners. The high number of audiovisual contents containing subtitles and the increase of multilingual audiovisual productions, required more attention on AST from the academic community, and this thesis aims to contribute to such an end by investigating the impact that the AST service and different strategies for their implementation have on the user's emotional activation and preference.

Conducting research on this underresearched topic finds a perfect match with the novelty of the methodology used in the experiments in this thesis. The study of user reception and its relation to the quality of the services offered to the audience has become a trend in the academic field (O'Hagan, 2016; Ramos, 2015). One of the most innovative approaches to user reception is the use of more objective measurements, such as the study of the body's physical reactions caused by psychological processes: psychophysiology (Ravaja, 2004). The application of psychophysiology to reception studies in the field of Media Accessibility (MA) is still a very recent practice that is slowly seeing progression and being established in the field (Fryer, 2013; Ramos, 2015). The reception of the services and any of their innovations must be tested in the research field in great detail so that, in practice, services can really improve and serve those people's specific needs.

The thesis, thus, finds place in a crossroads of disciplines, making it interdisciplinary. The combination of knowledge and methods from different areas allows for the localisation of new gaps in research, out of which thesis such as this one arises. The project found in this pages aims to respond to the challenge not only of exploring a topic greatly overlooked, but also at doing so by implementing methods scarcely applied in this kind of research and with this kind of stimuli.

1. Objectives and Hypotheses

This section presents the objectives of the thesis (main and secondary) and the hypotheses linked to these where applicable. The main objective of this PhD is the following:

To compare the levels of emotional activation produced by different AST delivery strategies in blind and partially sighted participants and analyse user preferences in relation to the delivery strategies.

Taking into account the previous works on the topic, this research aims to explore the characteristics of AST in depth by analysing the response of end-users to different types of AST delivery in filmic materials, particularly multilingual fictional films. In order to do so, two different strategies, voice-over effect and dubbing effect, are contrasted by testing their emotional impact on blind and partially sighted audiences and by gauging user preferences. Both psychophysiological measures and self-report instruments are used to that end. Linked to this main objective, the following hypotheses are put forward:

Hypothesis 1
AST with dubbing effect will be more emotionally activating than AST with a voice-over effect.

Hypothesis 2
Participants will show a preference for AST with a dubbing effect.

These hypotheses are grounded in the audiovisual culture of the participants, who are Spanish, and thus, highly exposed to dubbing products. A dubbing effect, where the original dialogue is removed, is predicted to be preferred over voice-over.

Based on the main objective, the following secondary objectives are put forward:

1. To define and describe AST through literature review and corpus analysis. More specifically our aims are:

1a) To carry out a literature survey that allow us to define the characteristics of this access service. In order to create a profile of the AST service, sources are sought in the academic world, the industry and governmental and standardising institutions.

1b) To undertake a qualitative analysis of audiovisual content in which AST is used.

This analysis takes into account voice-related features and multiple additional features, such as volume of both the original dialogues and their corresponding audio subtitles, isochrony, number of voice talents used, ways in which characters are assigned, and audibility of the original track. This analysis aims to provide a preliminary state of the art on current practices that sets the basis of the rest of the thesis. Linked to this aim, a third hypothesis is put forward:

Hypothesis 3
A high variability in terms of AST delivery is expected to be found.

2. To establish a theoretical framework by defining key concepts in this thesis such as voice-over/dubbing effect and emotional activation, and defining the theoretical models that guide the analysis.

Concerning voice-over/dubbing effect, they had been mentioned in previous works (Braun & Orero, 2010; ISO/IEC 20071-25, 2017) but only as a general categorisation. The formal and technical characteristics need to be well defined.

Regarding emotional activation, in accordance to the literature produced on cognitive processes and audiovisual media, it is profoundly interrelated with other concepts such as presence, flow and entertainment. It is important to clarify the concepts and terminology to be used in this thesis.

3. To compare the levels of emotional activation produced by AST with a dubbing effect and a voice-over effect on blind and partially sighted audiences to those of sighted audiences exposed to subtitled material.

Audio subtitling is provided as an alternative for those who cannot or do not want to read the subtitles. It is therefore relevant to test how the emotional activation produced by these alternative situations (AST with a dubbing or voice-over effect) compares to the situation in which a sighted viewer reads the subtitles. This is why results obtained in the experiments with blind and partially sighted participants, both with a dubbing and a voice-over effect, are compared to the levels of emotional activation provoked by subtitled clips on viewers who can actually benefit from all the visual elements of the product.

Hypothesis 4
Similar levels of emotional activation are expected to be found both in sighted audiences consuming subtitles and blind and partially sighted audiences consuming AST regardless of the effect.

2. Theoretical Framework

This project is born within the main framework of Audiovisual Translation (AVT) Studies, and most particularly within one of its branches: Media Accessibility. The object of the study, the AST service, embraces characteristics from a variety of audiovisual translation transfer modes and appears as a solution to the accessibility needs of some audiences.

Though few in number, the presence of AST in the literature tackles the service from different perspectives. From AVT itself, some authors have proposed different definitions and descriptions of their characteristics (Benecke, 2012; Braun & Orero, 2010; Remael, 2014). At the same time, from a more technical point of view, AST has been also analysed in some works (Mihkla et al., 2014; Nielsen & Bothe, 2008; Thrane, 2013) which are based on real experiences concerning the implementation of this service in real broadcasting.

On the other hand, although this project is situated within the branch of AVT Studies, and particularly, MA Studies, it draws knowledge from other disciplines, such as Psychology or Media Studies, from which theories on the effects of media on audiences are taken. Most specifically, the methodology carried out in the experimental section is set in the psychology branch of Psychophysiology.

For the purposes of our experiments, emotions, that can be read according to the dimensional (Lang, 1988) or discrete (Plutchik, 1980) approaches, have been conceptualised taking a hybrid approach, supported by recent authors (Harmon-Jones, Harmon-Jones, & Summerell, 2017).

This being said, this thesis acquires a transversal identity. This heterogeneity aims to produce a solid understanding of its central focus: the service of AST and its reception by end users. The theoretical background is found in Chapters 2 and 3.

3. Methodology

The thesis adopts a three-fold methodological approach: literature review, qualitative corpus analysis (Study 1) and, most importantly, user reception tests (Study 2).

Regarding the literature review, two main sources are used to characterise the service of audio subtitling and their multiple characteristics and applications: previous academic works and regulating documents.

Concerning the corpus analysis (Study 1), a group of six films which have been audio described and audio subtitled are analysed. The analysis is performed by means of a matrix of analysis that has been derived from a pilot analysis. This way, the characteristics of the AST service currently being applied in real commercial contents are gathered and described. More details are provided in Chapters 4 and 5.

The methodology used in user reception (Study 2) is the most innovative part of the thesis: the study of emotional activation by means of a psychophysiological approach together with self-report instruments. Such an approach is decidedly based on two referents: Fryer (2013) and Ramos (2015), who had previously applied such methodology to the study of AD. While Fryer worked on presence and its link to electrodermal activity and heart rate measurements, Ramos did the same with emotional activation but measuring heart rate. The methodology in this thesis relates both electrodermal activity and heart rate with emotional activation.

The experimental part of this thesis (Study 2) is divided in two main experiments whose results are later compared. Forty-two participants took part in each experiment.

For Experiment 1, three clips are selected from a Polish TV series (*War Girls*), which are professionally subtitled and then audio described and audio subtitled. The audio subtitling is made in two different versions: dubbing and voice-over. In order to test the impact of such two conditions in blind and visually impaired audiences, a self-report instrument (the T-SAM questionnaire) and a preference questionnaire are used together with psychophysiological measures.

Experiment 2 is performed with sighted participants. It presents the same methodology: the self-report instrument (T-SAM questionnaire) together with psychophysiological measures. They are used to test the impact of the subtitled version on participants. Further details on the methodology are provided in Chapter 6.

4. Thesis Structure

This thesis is structured in 8 chapters. **Chapter 1** defines the objectives and hypotheses and describes the theoretical and methodological framework for the thesis.

In **Chapter 2**, the notion of accessibility for people with vision and reading disabilities are presented (Section 2), particularly focused on the blind and the visually impaired since they are the participants of the experiments. The chapter follows with the definition of AST and their categorisations (Sections 3 and 4), previous research on AST (Section 5) and the topic of multilingualism (Section 6), since the experiments proposed in this thesis aim to be applicable to contents characterised by the existence of a main language and other secondary languages in the film.

Secondly, AST is analysed within the regulations, standards and guidelines across countries Europe and Western countries (Sections 8 and 9). The information mentioned directly about AST is scarce and reference to AST is sometimes found as part of documents related to AD. Therefore, AD is also considered in the analysis of regulations, standards and guidelines. The analysis of such documents is not reduced only to works issued by governmental institutions and private and public broadcasters, but also by some scholars and professionals.

Chapter 3, dealing with reception studies in AVT and MA (Section 1) aims to provide a theoretical basis for the methodology used in Study 2 (found in Chapters 4 and 5), emotional activation and psychophysiological measurement. The chapter describes the psychological processes that link the consumption of audiovisual material to entertainment and the experience of emotions (Sections 2 and 3).

The chapter defines emotion and the different approaches to emotion, which are described in detail as well as the hybrid model used in this thesis (Section 4). Finally, the focus is put on the methodology available for the measurement of emotions (Section 5) and the scarce number of studies within the study of AD and AST (Section 6).

Chapter 4 is devoted to the descriptive analysis of audio subtitled films. It refers to the method used for the description of a corpus of 6 multilingual audio subtitled films (Study 1). The results of the descriptive corpus analysis (Study 1) are discussed in **Chapter 5**, which summarises the main characteristics of the audio subtitles found in the selected corpus.

Chapter 6 deals with the methodology of the experimental part of the thesis (Study 2, including Experiments 1 and 2). Such methodology is based on the hybrid approach taken for the study of emotions and the combination of both self-reports instruments and psychophysiology. The chapter reports on stimuli selection, participants, materials, and experimental protocol.

In **Chapter 7**, discusses results from Study 2 (both Experiment 1 and Experiment 2) are presented. The chapter provides information on data treatment. The results obtained for the two conditions in Experiment 1 are compared, as well as the results from Experiment 1 and Experiment 2, that is, those performed with blind and visually impaired participants and sighted participants.

The thesis is finished in **Chapter 8**, where the conclusions are presented as a way of wrapping up the whole research process. Some remarks for future research are provided.

After the conclusions in Chapter 8, the **References** and **Filmography** of the thesis are found. Finally, the thesis is followed by the **Annexes**. Two kinds of annexes are provided: in paper, which are contained at the end of this document, and **Electronic Annexes**, distributed in electronic support.

The written **Annexes** contain, first of all, a table with a summary of the regulations, standards and guidelines that are contained in the literature survey of Chapter 2. Secondly, all the documents related to the experiments are presented, such as the information sheets and consent forms for the online validation of stimuli, the focus group for the validation of the self-report instrument and Experiments 1 and 2.

In the **Electronic Annexes** other kind of documents are found. In the folder “AST Scene Analysis Study 1”, an Excel with the transcription and all the information for each of the audio subtitled scenes of the 6 films conforming the corpus in the descriptive analysis is made available. In the folder “Raw Data”, all the psychophysiological raw data for both EDA and HR of all the 84 participants who took part in the experiment. The folder “Ethics Commission Authorisation” contains the documents certifying the approval of the ethics commission of the Universitat Autònoma de Barcelona on the experiments. The folder “Results Questionnaires” presents Excel files with the results of pre-questionnaires, Alexithymia questionnaires and preference questionnaires of Experiment 1. Finally, the folder “Stimuli” contains all the stimuli used in the development of the thesis, namely, those for the online validation and the final stimuli for Experiments 1 and 2.

Chapter 2

Theoretical Framework

Chapter 2
Theoretical Framework

This chapter provides a theoretical framework for this thesis. **Section 1** begins by contextualising the research within the framework of AVT, and most particularly within MA. **Section 2** moves to describe the main targeted end users and the needs AST can fulfil. **Section 3** provides a definition of audio subtitles, accounting for the terminological variation. Since AST is often provided in combination with AD, AD is very often referred to and links are made between both access services.

Based on a literature review, **Section 4** provides a categorisation of audio subtitles according to various parameters. The chapter then moves to **Section 5,** an overview of the main research works on AST. Next, the focus is put on the role of AST in multilingual films (**Section 6**), which are the ones under analysis in this thesis. The theoretical proposal by Sternberg (1981), applied to the subtitles for the deaf and hard-of-hearing (SDH) by Szarkowska, Zbikowska & Krejtz (2013), is adapted to AST as a possible framework for future analyses. In **Section 7,** this categorisation is then linked to the notions of dubbing and voice-over effects in the delivery of AST, which are the conditions that are tested in this research. The chapter concludes with an overview of regulations (**Section 8**), standards, and recommendations (**Section 9**) on AST and AD.

1. Audiovisual Translation and Media Accessibility

Nowadays society cannot be imagined without the strong presence of technology. Information is permanently flowing around us in various forms. We live in a world where communication has conquered the audiovisual channel as the way of dissemination. Audiovisual information is presented continuously throughout multiple platforms and perceived simultaneously by multiple senses: visual and audio. As a means of information transfer, translation has also been applied to this kind of contents. In academia, AVT was born as a new discipline in the second half of the 20^{th} century. The first publications on audiovisual translation date from the 60s with the appearance of three papers on cinematographic translation and dubbing (Caillé, 1960; Cary, 1960; Rowe, 1960). The discipline took shape and experienced a steady increase until the first conference was held in Stockholm in 1987[1]. From the 90s onwards and with the introduction of new technologies, its study has increased relevantly (Vercauteren, 2016, pp. 30-31) and is taught broadly across universities in the whole world (Orero, 2007c).

However, and just like most audiovisual products are translated into other languages, other ways of content adaptation are needed to cater for the different needs of users who cannot always access the content in its original form. Persons with hearing, vision, reading and cognitive disabilities cannot fully or partially access some of the inputs included in audiovisual materials. Therefore, to guarantee access for everybody, in the same way a language is translated into another form by means of dubbing, subtitling or voice-over (to name three audiovisual transfer modes), the audio or visual content is transferred into another mode by means of AD, AST or SDH to name a few. These access services are the object of research in MA, which can be defined as the

> set of theories, practices, services, technologies and instruments providing access to audiovisual media content for people who cannot, or cannot properly, access that content in its original form. Many of these instruments, of which the two most well-known are subtitling and audio description, stem from or involve audiovisual translation (AVT), because AVT is the field where MA has been developing as a research discipline for the last decade. (Greco, 2016, p. 23)

The study of MA, though recent, has produced a variety of new areas of research, ranging from descriptive to reception-based studies and encompassing various topics such as different AD styles (Walczak & Fryer, 2017), the reception of adapted materials (Fernández-Torné & Matamala, 2016), translation of AD scripts (Jankowska, 2015) or the segmentation and signalisation of SDH (Tsaousi, 2015).

[1] Dubbing and subtitling : European Broadcasting Union Conference, Stockholm 1987

Despite the existence of diverse user needs, two access services feature at the top of the list: SDH and AD. On the one hand, SDH

> offer a semantic account of what is broadcast in a particular program, but not only about how it is said (emphasis, voice tone, accent and foreign languages, voice noises) and who is saying it, but also about what is heard (music and ambient noise) and about the discursive elements appearing in the image (letters, legends, signs). [My translation][2] (Pereira, 2005, p. 162)

On the other hand, AD, targeted mainly at blind and partially sighted audiences, is considered as a form of mediation which is “cross-modal, involving essentially a 'translation' of visual images into verbal text” (Braun, 2008, p. 2).

Another access service which has not attracted so much attention is audio subtitling: it aims at the accessibility of people who cannot understand the language of the original and cannot make use of the translation provided by subtitles, be it because they have a vision disability or have difficulties reading text on screen.

2. The User: Accessibility for Visual and Reading Disabilities

AST can fulfil the different specific requirements of a broad range of users. When combined with AD, though, AST is commonly aimed at blind and partially sighted audiences. However, it can be argued that the application and delivery of AST can grant access to audiences with different kinds of sight loss, as the inability or trouble to read text from the screen is usually common to the entire collective.

Furthermore, it is important to highlight the existence of other audiences who, even if not blind, may not have full access to the visual content in an audiovisual content, particularly when it comes to written characters on screen. Elderly people or people with language and reading impairments, e.g. aphasia,

[2] “ofrece un recuento semántico de lo que se emite en el programa en cuestión, pero no solo de lo que se dice, cómo se dice (énfasis, tono de voz, acentos e idiomas extranjeros, ruidos de la voz) y quién lo dice sino también de lo que se oye (música y ruidos ambientales) y de los elementos discursivos que aparecen en la imagen (cartas, leyendas, carteles)” (Pereira, 2005, p. 162)

dyslexia (see Movietalk, 2014) or cognitive impairments such as intellectual disability or decreased concentration, are also likely to require the adaptation of the visual written input (Braun & Orero, 2010; Ljunglöf, Derbring, & Olsson, 2012). Indeed, the term “print disabled” may be applied in order to gather a broader group of people with specific reading needs. The UK Association for Accessible Formats describes a print disabled person as “anyone for whom visual, cognitive, or physical disability hinders the ability to read print. This includes all visual impairments, [and] dyslexia […]” (2012, p. 6). Print disability is understood mainly when considering physical print text (Accessible Instructional Materials For Maine Students With Print Disabilities, n.d.; Centre for Equitable Library Access, n.d.; Information Access for People with Print Disabilities, n.d.; Vision Australia, n.d.). However, in the previous sources it is suggested that in our digital society the concept of print disability could embrace other supports different from paper, since the difference is basically in the support of the written text.

AST has also been deemed useful for people who do not have any kind of visual, sensory or intellectual disability, such as children who have not yet entirely developed their reading skills, learners of a second language (Mihkla et al., 2014; Mihkla et al., 2013) or as mentioned before, the elderly or persons with reading difficulties (low literacy, etc.). Even if it is difficult to establish exact numbers of people with specific visual or reading needs – since the definition of disability varies politically and administratively from country to country - what seems common to all demographic analyses is the relationship between capability and age. Therefore, we have to keep in mind that, for instance, more than 70 million people are currently aged 60 or above in the EU. Many of these citizens will “experience […] cognitive, hearing and sight problems in later life” (Stallard, 2003, p. 5). Due to the population ageing the number of people requiring such services will increase.

3. Defining AST

After understanding the audiences that may potentially need audio subtitled contents, it has to be mentioned that AST is intended primarily for the blind and the partially sighted in its combination with AD. In fact, AST is often found as a part of the AD track in commercial accessible films and other audiovisual contents. Due to the relationship between the two services, this section briefly defines AD (Section 3.1.) and then moves to a definition of AST (Section 3.2.).

3.1. AD: Transforming Images into Words

In his categorisation of forms of translating the verbal sign, Jakobson (2000 /1959) proposed the concept of intersemiotic translation when a message is translated into a different type of sign. AD is defined as an intersemiotic and intermodal (Braun, 2008) audiovisual transfer mode that involves the translation of visual images into verbal descriptions (Maszerowska, Matamala, & Orero, 2014). It takes the form of “an additional narration that fits in the silences between dialogues and describes action, body language, facial expressions and anything that will help people with visual impairment follow what is happening on screen or stage” (Díaz Cintas & Anderman, 2009, p. 8). According to Benecke & Dosch (2004) and Pujol & Orero (2007), AD can be traced to ancient times when in the Medieval Ages, pilgrims needed oral explanations to read and enjoy the visual narrative in the glass windows of churches and cathedrals. The academic study of AD within the framework of AVT and MA has drawn an increasing interest in recent decades (Igareda & Matamala, 2012; Kruger & Orero, 2010; Maszerowska et al., 2014, Matamala & Orero, 2016) and has become available in an increasing number of countries in the last 20 years (Fryer & Freeman, 2013; Orero, 2007c).

The origins of AD as an accessibility service are found in the 70s, when Margaret Pfanstiehl established the Metropolitan Washington Ear, Inc., a non-profit organisation and radio station devoted to enriching the lives of the blind and people with other kinds of impairments (Pfanstiehl, 2016). However, in Spain, other kinds of emissions, which were not initially created as accessibility services, were broadcast. Already in the 1940s, after the civil war, the first films were “described” on the radio (Jankowska, 2015; Orero, 2007b). After that, more countries started offering these services, both from associations of blind people and from governments and broadcasters (Orero, 2007c).

However, less has been said about the access service that concerns this thesis: AST. De Jong, in his contribution to one European Broadcasting Union Technical Report, mentions that “audio description is a much more elaborated service than spoken subtitling, which is limited to ‘reading subtitles’” (2004, p. 5). This assertion, despite being true to some extent, should be revisited.

3.2. AST: Converting Written Text into Sound

AST is placed somewhere between the accessibility of audiovisual contents for specific needs and the more canonical conception of the translation of text. The conception is based on the translation proper (interlingual translation) proposed by Jakobson (2000 [1959]). It is based on the transference of a message in a source language into a target language. Accessibility and translation mean that a source language is translated into a target language in the written subtitles and then delivered orally in a different sign.

Audio subtitles or spoken subtitles (Ljunglöf et al., 2012; Verboom, Crombie, Dijk, & Theunisz, 2002) have been defined in a number of works so far. Most of the literature feature the idea of “spoken or aural rendering”, such is the case of Nielsen & Bothe (2008), Miesenberger, Klaus & Zagler (2002, p. 296), Fryer (2013, p. 194), Szarkowska & Mączyńska (2011, p. 24), Orero (2011, p. 237), the International Organization for Standardization & International Electrotechnical Commission (2015) or the International Telecommunication Union (Focus Group on Audiovisual Media Accessibility) (2013, p. 3). Some definitions of AST are applied specially to computer systems that allow for its delivery (by speech synthesis). These definitions state that such AST is enabled by computers to “read subtitles of movies and TV shows aloud” (Ljunglöf, Derbring, & Olsson, 2012, p. 1) or that they are created automatically in real time (Verboom et al., 2002). Thrane (2013, p. 12) by taking the definition provided in the webpage of the Centre for Excellence in Universal Design (2013), specifies that such “reading aloud” in a “spoken voice” is issued for “viewers with vision or reading impairments”.

Harrouet (2016, p. 8), in a more holistic way, provides a more etymologically-driven definition stating that the term comes from two other transfer modes in AVT: subtitling and AD. AST, applied particularly to cinema and other fiction genres, is described as a “hybrid variant” of subtitles that transforms the dialogues of the film. The dialogues, which are “written to be spoken, then summarized and rewritten to be read as subtitles by the audience, now have to be read out loud again as audio subtitles” (Remael, 2014). Such hybridism places “this modality of audiovisual localisation […] at the interface between subtitling, audio description and voice-over (Braun & Orero, 2010, p. 1).

However, one of the definitions that can be deemed to embrace best the heterogeneity of the application and form of this service is the one provided by Reviers & Remael (2015, p. 52):

> AST can therefore be defined as the aurally rendered and recorded version of the subtitles with a film. This spoken version of the subtitles is mixed with the original sound track. AST is usually read, sometimes acted out, by one or more voice actors. Sometimes it is produced by text-to-speech software. The subtitle text is often delivered almost literally, but it can be rewritten to varying degrees, and in addition, the recording method also varies. Usually, AST is recorded as a form of voice-over, which means that the original dialogues can be heard briefly before the translation starts. Sometimes it is recorded in a semi-dubbed form, which means that the original dialogues are substituted by a form of dubbing that is not necessarily entirely lip-sync, that is, synchronous with the lip movement of the speaker.

All definitions refer to the audio rendering of interlingual subtitles, that is, in a language different from the original. This is especially relevant in subtitling countries, but also very relevant in other territories. Even in traditionally dubbing countries such as Spain, fully subtitled contents are increasingly accepted (Matamala, Perego, & Bottiroli, 2017) and interlingual subtitles are more present in multilingual audiovisual contents. Such products include excerpts in secondary languages, that, even if the main language is dubbed, can be translated by means of written subtitles. Multilingualism is treated further on in this chapter (see Section 6).

4. Categorising Audio Subtitles

Audio subtitles can be classified according to different parameters, such as the combination of AST with the AD, whether or not audio subtitles are autonomous and are found independently or inside the AD track (integrated); the mix, that is, when and how AST is integrated with the rest of tracks; and the voice used for their delivery: a synthetic or a human voice. The next sub-sections delve into these categorisations.

4.1. Combination: Autonomous and Integrated

Audio subtitles can be offered independently or integrated in the AD track. The terms used in this thesis are autonomous audio subtitles for the first type and integrated audio subtitles for the second type.

Autonomous audio subtitles are commonly found in television in countries with strong subtitling traditions, such as the Netherlands, Sweden, Denmark or Finland. They are made available to final users by means of systems such as Text-to-Speech (TTS) and Optical Character Recognition (OCR). In the Netherlands, the FSB (Netherlands Federation of the Blind and Partially Sighted), the NOS (Public Broadcasting Company) and the FNB (Federation of Dutch Libraries for the Blind) collaborated in the Spoken Subtitles Initiative to implement this system on a regular basis since 2002 (Klaus Miesenberger, Klaus, & Zagler, 2002, pp. 295-302; Verboom et al., 2002). In Sweden, the National Swedish Television Station (SVT) started broadcasting AST in 2005 (Ljunglöf et al., 2012; Nielsen & Bothe, 2008) and some systems were developed for the voicing of subtitle files also for computers (Derbring, Ljunglöf, & Olsson, 2009). This project was not continued and currently new TTS software programs through new supports, such as smartphones, are used to read subtitles aloud.

The Danish public channel, DR1, launched in 2012 a pilot project, DR1Syn, a twin channel in which every time a foreign language was spoken AST was made available (Thrane, 2013). The Danish system is no longer used since the change from analogue to digital technology required an update in the spoken subtitle service. Similarly, in 2008 the Czech Republic also started the ELJABR project (Elimination of the Language Barriers Faced by the Handicapped Watchers of the Czech Television) aimed at the implementation of AST in Czech TV (Hanzliček, Matoušek, & Tihelka, 2008). Although some more development was made to the system (Matousek & Vit, 2012), the system is no longer used and the method is being updated to suit the new standard.

Integrated AST are merged into the AD track. They are the usual option for the production of filmic AD, as it can be seen in regular commercialised DVDs. In research, examples are found, such in the case of Remael (2012) on Dutch multilingual films or Benecke (2012) on German dubbed films. While autonomous AST is normally delivered with synthetic voices, AST integrated in the AD track for cinema are normally recorded in a studio by voice talents at the same time as the AD.

4.2. Voices: Synthetic and Human

Audio subtitles can also be categorised depending on the type of voice used for their delivery: human and synthetic voices. The type of voice has already been studied in the implementation of AD in terms of acceptance and user preference (Fernández-Torné & Matamala, 2016; Szarkowska, 2011; Szarkowska & Jankowska, 2012; Szarkowska & Walczak, 2012; Walczak, 2010; Walczak & Fryer, 2018).

Fernández-Torné & Matamala (2016) mention that synthetic voices in AD are generally accepted and particularly if they are used in short audiovisual productions. The technological advances in the quality and implementation of synthetic voices provide a useful way to offer more autonomy to audiences in need of them. However, there is still preference for human voices in fiction contents. They are easier to follow, especially in films where audiences are supposed to feel changes in feeling and particularly for children and young people (Holsanova et al., 2015). However, most part of studies agree on the fact that synthetic voices, despite providing autonomy, easy availability to users and being generally accepted by audiences, present some limitations in terms of quality (Holsanova, Hildén, Salmson, & Tundell, 2015; Nielsen & Bothe, 2008) and character recognition (Derbring et al., 2009; Ljunglöf et al., 2012). Furthermore, the acceptance of such synthetic voices depends also on the genre in which we use them. For instance, in Thrane’s study (2013) the use of synthetic voices for AST was very well rated by end-users when it came to news or documentary excerpts.

On the other hand, human voices are commonly used in AST for cinema as a part of the AD track in three mains scenarios: (a) films showing any written on-screen text (other than the written subtitles for the dialogues), (b) films presenting written subtitles for other kinds of language, such as sign or tactile language (Orero, 2011), and (c) films partially in another language unknown to the audience (Benecke, 2012; Braun & Orero, 2010; Remael, 2012; Reviers & Remael, 2015). Even in non-screen artistic productions, such as opera, live human-voiced audio surtitles proved to be quite effective and accepted by the audience in one of the first user experiments on live audio surtitles in opera performances (Orero, 2007a). *Roberto Devereux* by Donizetti, in a minimalist production where not a lot of AD was needed, was presented with human-voiced AST for visually impaired audiences.

4.3. Mix: Broadcast-Mix and Receiver-Mix

When considering how the audio-rendering of the subtitles is made available to final users, the main distinction is based on the way the two channels or tracks (that of the original soundtrack and the AST track) are mixed: broadcast-mixed (Media Access Australia, 2012), also called pre-mixed (ISO/IEC 20071-25, 2017), or receiver-mixed.

The distinction is presented by Ofcom (Tanton, 2015) in relation to AD but can be adapted to AST. In the first case, the AST track is mixed with the original track before reaching the final user. In other words, the company or author of the DVD, film or TV programme creates a second track in which the original track of the audiovisual product is mixed with the AST. In the latter, the merging of both tracks takes place at the end user's device(s) or receptor(s). It "involves mixing decoded programme sound and decoded description in the end-user-equipment" (Tanton, 2015, p. 1). The AST (and/or AD) track is sent separately and can be played along the original track with signals that lower the sound level of the original so the AST stream can be heard (see ISO & IEC, 2017, p. 8; Media Access Australia, 2012). Receiver-mixed AST allows the user to control the volume of the AST and can be fed to the headphones without the rest of the room hearing the AST (Media Access Australia, 2012).

5. Research on AST

Research on AST is scarce compared to the body of research on AD, which tackle issues such as the form and structure of AD and its possible applications, for example in opera (Cabeza-Cáceres & Matamala, 2008a; Corral & Lladó, 2011), the audio introductions (AI) for cinema (Romero-Fresco & Fryer, 2014a) or the translation of AD scripts (Jankowska, 2015). Furthermore, the applications and characteristics of AD have been put to test in a variety of reception studies (see Chapter 3).

One of the first studies on AST focused on the access to the surtitles for opera (Orero, 2007a). The application of the AST of surtitles by means of a human voice was implemented at the Gran Teatre del Liceu, in Barcelona. Orero took as inspiration the project developed in the Netherlands by Verboom et al. (2002). In such project, though, subtitles were made accessible with a TTS system by means of synthetic voices and it was done for regular TV broadcast.

Some years later, Braun & Orero (2010) analysed the first 10 minutes of 14 audio described and audio subtitled films[3]. They identified that written subtitles normally condense information and that they rely on the visual input as an anchorage, which can pose a problem for people who cannot access the visuals (pp. 175, 176). To introduce the audio subtitles, different strategies are proposed in Braun & Orero's analysis (2010):

1. Creating a change in the intonation when reading aloud the subtitle,
2. Preceding the subtitle/caption delivery by "a subtitle/caption reads",
3. Delivering directly in a quote the utterance within the AD content,
4. Providing the name of the speaker in front of the corresponding subtitle, and, in a more indirect way,
5. Including a description of the situation, such as the direction towards a character's gaze is oriented to.

Benecke (2012) analysed different situations where AST is used in dubbed films. He focused on secondary language(s) which are relevant to the plot, because the main language is fully dubbed. Benecke observed that in *The Piano* (Campion, 1993), sign language changes rapidly to spoken language and

[3] *Be with Me* (Khoo, 2005), *Borat* (Charles, 2006), *Brick Lane* (Gavron, 2007), *Bride and Prejudice* (Chadha, 2004), *East is East* (O'Donnell, 1999), *Everything is Illuminated* (Schreiber, 2005), *Hero* (Yimou, 2004), *Kung Fu Hustle* (Chow, 2004), *Letters from Iwo Jima* (Eastwood, 2006), *Munich* (Spielberg, 2005), *The Passion of the Christ* (Gibson, 2004), *The Science of Sleep* (Gondry, 2006), *Syriana* (Gaham, 2005) and *Volver* (Almodóvar, 2006).

the AST results confusing. In the German *Am Limit* (Dunquart, 2007) the speakers of English utterances are mentioned by the audio describer. AST in multilingual films such as the polyglot *Inglorious Basterds* (Tarantino, 2009) and *Babel* (González Iñárritu, 2006) different voices are used to create the AST of various characters speaking in a different language other than the main language of the film. The audio subtitling of songs maintained in the original language is exemplified in films such as *Volver* (Almodóvar, 2006), where the same female audio describer provides AST of the song and describes the action or *Hable con ella* (Almodóvar, 2002), where a female voice voices the subtitles for the song (sung by a woman) whereas the AD is provided by a male voice. Despite being monolingual films, these cases serve as example of a trend in foreign films that are dubbed and which include original songs, left in the original language and subtitled with written text.

Remael (2012) analysed multilingualism in the Dutch/Flemish context. The article focuses on two Dutch films, *Oorlogswinter* (Koolhoven, 2008) and *Tirza* (van den Berg, 2010). She highlighted the fact that some of the Dutch actors were able to speak the foreign languages to be subtitled (for sighted audiences). This situation allowed the same actors to voice the audio subtitles in the film. It remains unclear whether this strategy results in better coherence and general comprehension of the dialogues. Of great interest is that these films were audio described and audio subtitled by professionals taking into account Braun & Orero's (2008) challenges mentioned before. Remael (2012) brought up four main topics: the role of the AD when delivering the AST, the manipulation of text between written subtitle and AST, the audibility of the original lines, and the importance of intonation.

In the first film, *Oorlogswinter,* the AD provides a lot of information about the position and movements of the characters and the description of facial expressions, facilitating character recognition when audio subtitles are heard. In this film AST and written subtitles are practically identical, although some discussion about the condensed written nature of standard subtitles might arise around this point. Remael highlighted the importance of intonation when recognising the character talking. In the film analysed professionals made a great effort to render the performance of the actors, speaking in a foreign language, into the AST. Finally, she mentioned the duality of effects in AST, as in Orero & Braun (2010): dubbing and voice-over effects. Whereas *Oorlogswinter* is said to present a voice-over effect, since AST are delayed in relation to the original dialogues, the intention in *Tirza* was to cover up the original dialogues in an almost complete dubbing effect. However, in both films original dialogues are audible in different volumes. This practice is based on the commentary by the sound technician of *Oorlogswinter*, where in some scenes the original remains present almost as an echo to indicate that a foreign language is being spoken. The analysis of these two strategies are later explained in deep.

Matamala (2014), in an article which aimed to embrace the adaptation of all types of text on-screen, also analysed the challenges that may be posed by the audio description of the text appearing on screen in *Inglorious Basterds.* Apart from her analysis of the possibilities for the adaptation of all written text, such as newspaper clippings and maps, she treated with particular interest the multilingualism in the film. In fact, English, German, French and Italian are spoken and apart from English, which is the predominant language, the rest of languages are subtitled in the original version. When discussing AST, Matamala (2014) pointed at the possibility of not voicing those very short excerpts featuring a foreign language, which are not even subtitled for sighted audiences. She provided a comprehensive summary of the potential strategies for the delivery of AST.

The complexity of this access service is also described in Reviers & Remael (2015) in terms of what they called “multimodal cohesion” and how this affected the coherence perceived by the audience. In such study the multilingual film *Oorlogswinter* (Koolhoven, 2008), already used in Remael (2012), and *Süskind* (van den Berg, 2012) were surveyed. The authors conclude that the audio described film maintains its coherence thanks to the links between AST and AD allowing the audience with specific needs to recreate the same experience in terms of narrative coherence. Although the production of AST for both monolingual and multilingual films can be seen in parallel, the thesis focuses on the latter. Multilingualism is a topic which has been considered by other scholars when studying AVT transfer modes, such as dubbing or subtitling. It is interesting extrapolating such information and applying such knowledge to the production of AST.

6. AST and Multilingualism

"Directors are incorporating foreign languages into their scripts in order to add a touch of authenticity to scenes set in foreign environments, thus allowing the audience to immerse into 'the foreign'", state Szarkowska et al. (2013). The multilingualism in some films is an attempt to represent the linguistic reality. See, for instance, De Higes-Andino's analysis (2014) of multilingualism in the film *It's a Free World…* (Loach, 2007), which shows a British society characterised by migration and diaspora. When a film portraying different languages is to be distributed and showed in other countries, the problem of translation and accessibility arises.

Some scholars within AVT have focused their attention in the treatment of multilingualism. Taking an overall approach, De Higes-Andino (2014) elaborated an analysis tool of translation strategies according to different transfer modes in the translation of multilingualism and applied it in a case study. Other studies have compared how different transfer modes, such as subtitling and voice-over, convey multilingualism and its impact in terms of comprehension (Sepielak, 2016b).

The strategies shown in multilingual films exported to dubbing countries were the object of research by Heiss (2004) and Voellmer & Zabalbeascoa (2014), who observed that multilingualism is maintained as a regular practice and that when secondary languages are portrayed written subtitles are normally used. Other strategies include the representation of a strong accent in the dubbing of secondary foreign languages (Voellmer & Zabalbeascoa, 2014). Bartoll (2006), in turn, focused on the strategies to be applied to subtitling in order to convey the multilingual reality of a film. More interestingly, he drew practices from MA, such as the use of extra information or different colouring of the subtitles normally related to the SDH.

Multilingual films pose important challenges not only in terms of translation but also in terms of accessibility for people who have no or limited access to the visuals (Benecke, 2012; Braun & Orero, 2010; Espasa & Corrius, 2017; Harrouet, 2016; Remael, 2012; Reviers & Remael, 2015). It has been shown that AST is the most usual way of portraying and handling the multiplicity of languages in an audiovisual product. The next section discusses the treatment of multilingualism through this access service.

6.1. Revealing Multilingualism with AST

In the beginning, only independent films were aiming to show the rich complexity of multilingual realities in human society (Bleichenbacher, 2008, p. 21) but this practice has steadily become a trend in more Hollywood mainstream cinema production (Braun & Orero, 2010). In countries with a subtitling tradition, where all foreign languages are subtitled, AST seems the most reasonable option for the aforementioned audiences with specific needs. In dubbing countries, such as Spain, foreign languages in a film or any other audiovisual content are normally replaced completely with dialogues translated into the target language. However, secondary languages are often subtitled, generating a partially subtitled film.

The way multilingualism is revealed to the audience was approached by Szarkowska et al. (2013) in the production of SDH. This kind of subtitles, similarly to AST, provide information that can be accessed only by using one sensory channel, in this case, sight. By displaying the SDH in different forms, using different labels and information, Szarkowska et al. (2013) created a scale from less to more foreign-revealing. Their categorisation is a reinterpretation of Sternberg's representation of polylingualism (1981), which is adapted here to AST. Even if the following parallelism between SDH and AST in multilingual films is mainly of theoretical ilk, it aims to show the different approaches that could be taken when making accessible multilingual dialogues through the use of AST. Sternberg's original representation includes six levels for rendering multilingualism in the translation of text:

1. **Vehicular matching:** which implies the presence of the contents in the original foreign language without further information, requiring a particular knowledge from both the producer and the receiver of the text;

2. **Selective reproduction:** which is based on the "intermittent quotation" of the original foreign language, that is, providing the receiver with hints that

the text is originally delivered in a foreign language, thus showing the multilingual discourse;

3. **Verbal transposition:** which presents a form of imitation of some of the features of the foreign language in the language understood by the receiver, such as phonetic characteristics or syntactic structures.

4. **Conceptual reflection:** which is more concerned with the socio-cultural rather than the linguistic, as an example, Sternberg mentions the description of a watch by the Lilliputians in the oeuvre *Gulliver's travels* when they do not know the object nor the name of it (the watch) in their vocabulary.

5. **Explicit attribution:** which appears as a "direct statement of the reporter's […] part concerning the language […] in which the reported speech was originally made" (p. 231) (e.g. *He says in Polish*).

6. **Homogenising convention:** which dismisses all variations in languages that may appear in the text by producing a monolingual discourse.

The scale was then applied by Szarkowska et al. (2013) to the portrayal of multilingualism through SDH. They combined Sternberg's categorisation with other strategies intrinsically belonging to SDH, such as colour coding, and differentiated 5 levels:

1. **Vehicular matching:** the direct transcription of the foreign language spoken.

2. **Translation and explicit attribution:** translating foreign utterance and indicating the language spoken.

3. **Translation and colour-coding:** includes a translation and colour-coding it without revealing the language spoken.

4. **Explicit attribution:** consists of just telling the receiver that a foreign language is being used.

5. **Linguistic homogenisation:** implies the translation of the utterance without indicating that it is a translation.

They excluded what Stenberg describes as “conceptual reflection” as this is related to the characteristics of the textual translation (semantics) and socio-cultural factors. It is not, then, related to a concept that could be rendered as a strategy for the adaptation of multilingualism through SDH. The strategies they proposed were tested with final users. As a general outcome, they concluded that deaf and hard-of-hearing users preferred more rather than less information about the language(s) spoken in the dialogue.

Some parallelism can be drawn between the way SDH and AST portray information about multilingualism. This can be written in brackets in the SDH and verbally informed in both the AST or AD. On the other hand, information can be provided by means of strategies such as colour-coding for SDH or the audible presence of the original in AST, a feature related to the dubbing and voice-over effect in the way audio subtitles are delivered, explained in detail in the next section. The categories described by Sternberg have to be adapted to suit the differences in the characteristics of the texts (written and audible) and the genre (literature and film). Despite the differences in the adaptation of Sternberg’s representation, a parallel scale from the more multilingualism-revealing strategy to the more homogenising strategy is proposed for AST:

1. **Vehicular matching:** presents the original soundtrack without any other information (without AST or any information provided in the AD track. The comprehension of the content and the language spoken depends only on the receiver’s linguistic knowledge.

2. **Selective reproduction:** implements AST with a voice-over effect in the target language. This allows the user to understand that this is a translation. The foreign language is heard in the background without providing further information on the language spoken.

3. **Selective reproduction + language information:** AST is provided with a voice-over effect, similar to selective reproduction, and the audio describer provides information on the language spoken.

4. **Verbal transposition:** AST goes beyond the translation by imitating and reproducing some of the patterns of the original foreign language spoken. It could be used with both a dubbing or a voice-over effect, since the reproduction of the features concern phonetic, semantic and syntactic structures found in the foreign language that is being portrayed.

5. **Explicit attribution:** is acquired by providing an AST with a dubbing effect after the language to be spoken is announced; this produces a result closer to the complete homogenisation of the contents.

6. **Homogenising convention:** is simply based on the complete translation of the original line by means of full dubbing, erasing any traces of the foreign language and creating a monolingual discourse.

The different treatment of AST in terms of technical variances, such as audibility or synchronisation result in what is here referred to as “effect”. From previous descriptive studies (Braun & Orero, 2010), and some standards (ISO/IEC 20071-25, 2017), the two main effects identified in the delivery of AST are dubbing and voice-over effect, which are presented in the next section.

7. Dubbing and Voice-over Effect

In order to describe the features of these two effects, a definition of the two AVT transfer modes from which they draw their characteristics and names will be provided: voice-over and dubbing.

The transfer mode of voice-over (Franco, Matamala, & Orero, 2010; Matamala, 2018) is based on the superimposed spoken rendering of the dialogues or monologues in another language on a translated voice. It is delivered simultaneously with original speech length but it does not account for lip synchronisation and starts sometime after or before the original. It may reproduce to some extent mimetic features (accent, age, emotion, gender, intonation, orality markers and stress). The volume of the original is brought to a low level and is still heard in the background. It allows the audience to hear the original and the translated speech. When the translation finishes, the sound of the original is restored to normal (Díaz Cintas & Orero, 2006, p. 477). The overlapping of voices is what differentiates the term voice-over in translation studies and in film studies (Matamala, 2018).

On the other hand, dubbing "consists of replacing the original track of a film's (or any audiovisual text) source language dialogues with another track on which translated dialogues have been recorded in the target language" (Chaume, 2012: p. 1). Chaume enumerates different characteristics of this type of revoicing (pp. 14-21), as he puts it. Dubbing has to achieve an acceptable lip-sync, synchronisation being one of the key aspects of dubbing; clear sound quality, in which the "dialogues from the original version must never be heard" and "the volume of the voices is also higher than in normal speech" (p. 18); and acting, that "includes the performance and dramatization of the dialogues" (p. 19) on behalf of the dubbing actors and actresses.

When applied to the voicing of AST, the two effects have been only mentioned in three documents. Some parts of the descriptions provided for both effects mirror some of the characteristics of the two transfer modes described above, thus allowing for a parallelism. The term "effect" implies resemblance, thus, AST are not delivered strictly using a transfer mode but they do so in a way that falls closer to one or the other depending on their features.

The distinction between both effects is introduced by Braun & Orero (2010) in the following terms:

- Voice-over effect: "achieved when the original dialogue can be heard while the audio subtitle is superimposed." (p. 181)

- Dubbing effect: "achieved when the ASTs and the original dialogue are delivered synchronously without the original voices being heard." (p. 181)

The ISO/IEC 20071-25 standard (2017) also mentions the dichotomy between both effects:

> when spoken captions/subtitles are delivered in synchronization with the captions/subtitles on screen, the original dialogue might be heard on the background or not. When the original dialogue is not heard, the final effect is similar to dubbing. When the original dialogue is heard underneath, a voice-over effect is achieved. (p. 12)

Finally, in the section of the recommendations on AST in the ADLAB project (Remael, 2014), further information is given about the two effects too. They mention that “the two most common ways for recording AST are voice-over and *a form of* dubbing”. When delivering AST with a voice-over effect, the AST starts “seconds after the original”, allowing it to be heard “in the background” (n.p.). This effect is said to help the audience to identify the speaker.

These documents also provide information on the prosody in the two effects. Braun & Orero (2010) talk about “intonation patterns” (p. 182) when they discuss an audio subtitled being “acted out” (p. 182). The ISO/IEC 20071-25 (2017) mention that “spoken captions/subtitles can be delivered in an acted or a non-acted way” (p. 14) and that “voices with different accents can also be used” (p. 14). Furthermore, it is highlighted that “changing the prosody” (p. 15) can facilitate distinguishing between when a text is being read or when the character talking. Prosody is mentioned in the definition of the two techniques. For the definition of voice-over, the standard mentions that “mimetic features” can be reproduced “to a certain extent” (p. 15). In dubbing, as a technique that involves replacing the entire original dialogues, prosody is imitated, in the sense that the dialogues are acted-out, as put by Braun & Orero (2010). In the recommendations of the ADLAB project the dubbing effect is related to “more “acting”. It is important to highlight the fact that the word “acting” is written between quotation marks probably implying that the term is not clearly defined.

Actually, there is no neutral delivery of AST, and in both effects the voice talent acts, but in a different way. In this thesis, the notion of non-acted has been related to that of a reading aloud prosody, closer to the prosody present in documentaries, as opposed to an acted voice, that is more related to the recreation of a fake orality found in dubbing. On the one hand, read aloud speech is different to oral natural speech, which contains usually more pauses than those suggested by punctuation, especially when long sentences are uttered (Schwanenflugel, Hamilton, Wisenbaker, Ruhn, & Stahl, 2009). The oral reader "must abstract prosodic features to a great extent while reading aloud" (p. 3). Indeed, Schreiber (1987) states that the explicit presence of prosodic features is the main difference between oral speech and reading, and such features make oral speech easier to comprehend. On the other hand, the artificiality of dubbing makes this specific way of speech to be referred to as "dubbese", which actually tries to imitate a fake "natural" orality (Baños & Chaume, 2009; Sánchez Mompeán, 2016).

Another key element in the distinction of the effects is synchronisation. Chaume (2004) defines synchronisation in dubbing as "the process of recording a translation in any given target language in a dubbing studio, matching the translation with the screen actors' body movements and articulatory movements" (p. 39). Three types of synchronisation are presented by the author:

- Phonetic or lip-synchrony
- Kinetic synchrony or body movement synchrony
- Isochrony or synchrony between utterances and pauses

The third type, "isochrony or synchrony between utterances and pauses", is a key element when making the distinction between dubbing and voice-over effect. Whereas synchronisation in dubbing aims at providing a perfect isochrony of both tracks, voice-over's synchronisation implies that the translation "should fit in the time available […] in such a way that the beginning and end of the original utterance are audible" (Sepielak, 2016a, p. 1060). In this regard, the ADLAB document explains that the voice-over effect implies a delay in the display of the AST regarding the original line. In conclusion, the dichotomy separates:

1. **Voice-over** effect: in which the AST is displayed a short time after the original, which can be heard underneath. The voice used can be acted to some extent but its prosodic features are related to reading-aloud, more faithful to the written code.

2. **Dubbing** effect: in which the AST is displayed in synchrony with the original line (isochrony), which cannot be heard. The voice used is acted, in the sense that it replicates the emotive lead in the dialogue, similar to the standardised prefabricated natural orality attributed to dubbing.

In multilingual films, such effects in the delivery of human-voiced AST for fiction AV products are combined with the information provided by the AD in order to reveal more or less the linguistic plurality in the scene. As presented in the readaptation of Steinberg's representation (1981), a voice-over effect in general is considered as more multilingualism-revealing, allowing the original language to be heard in the background. If provided, the information given by the AD in terms of the language is spoken in the original will place the AST in a different position of the step along the multilingualism-revealing scale.

Dubbing, on the other hand, is considered to be more domesticating, that is, hiding multilingualism, as does not allow the original language to be heard during the AST or before it. Again, if the AD mentions more or less information about the conversational situation (language spoken) such AST with dubbing effect would have a higher or lower position in the categorisation.

8. Regulating Media Accessibility: AST and AD

This section is composed of the analysis of the regulations regarding MA in some countries. The vision here exposed is limited. It does not include other territories such as African, Asian or South and Central American Countries. It focuses on European countries and main Western countries in terms of cinema production, i.e. those contributing relevantly to Western cinema and audiovisual production in general, which are English-speaking countries such as the United States, Canada or Australia. The analysis is based on the regulations survey conducted by the Agency for Fundamental Rights of the European Union

(Agency for Fundamental Rights, 2014), and complemented by online databases on media and MA regulations, such as the section on Media Legislation of the European Platform of Regulatory Authorities (European Platform of Regulatory Authorities, 2018) or the Accessometer of the Media Accessibility Platform (Media Accessibility Platform, n.d.).

Some countries mention MA in some of their acts but do not provide any kind of further information about how they implement it. They only encourage or recommend broadcasting companies to steadily increase MA services. This is the case of countries such as Austria (National Action Plan on Disability 2012-2020, 2012), Croatia (Act on Croatian Radio-television, 2012; Electronic Media Act, 2013), Malta (Broadcasting Act, 2015), Portugal (Lei da Televisao e Serviços Audiovisuais a Pedido, 2011), Romania (The Audiovisual Law, 2002), Luxembourg (Law on Electronic Media, 2015) or Latvia (Electronic Mass Media Law, 2010) or Lithuania (Law on the Provision of Information to the People, 2015). Germany aims to also increase the offer of accessible programs in television (Interstate Broadcasting Treaty, 2017). Besides, Slovenia, in its Audiovisual Media Services Directive (2011), states the need for more accessible services as a recommendation. Similar is the situation in Italy, where even if the contract (Ministero dello Sviluppo Economico, 2018) of the government with the main broadcaster, the RAI, states that the broadcasting channel guarantees access to people with sensory or cognitive disabilities, no further information is given regarding the amount of accessible programs. In Canada, its legislation do not impose any kind of figure regarding the amount of accessible programs (Broadcasting Act, 2018).

Countries such as Estonia (Media Services Act, 2014), even if they include a vast range of services such as subtitles, sign language, translation, separate audio channels, etc., do not offer any kind of compulsory service to be implemented by their broadcasters. Denmark (Danish Ministry of Culture, 2018), in its contract between the government and the main broadcasting institution, proposes better knowledge of accessibility services and invites a regular dialogue with relevant associations of persons with disabilities.

The last two paragraphs above could be put in parallel with European regulations (European Parliament, 2010), in which there is no information on the amount of programming that should be made accessible and the broadcasters that should offer it. The Audiovisual Media Services Directive, for example, makes reference to the need of making all media content accessible for people with a variety of specific needs and encourages media providers to make contents accessible (Audiovisual Media Services Directive, 2010). The contents of the forthcoming European Accessibility Act are being discussed (European Parliament, 2016), although the official document is still to be released. This document will then produce changes in the contents of the regulations of many countries.

Fortunately, there is a group of countries in which MA is made a more plural and mandatory issue by state institutions. The region of Catalonia, which stands as one of the regions offering most audiovisual accessibility services in Spain, presents not only an agreement between the government and the Audiovisual Council of Catalonia (Consell de l'Audiovisual de Catalunya, 2013) but also periodical reports of the accessible programs broadcast in the regional television (TV3), the last of them issued in 2017 (Corporació Catalana de Mitjans Audiovisuals, 2016). This last report shows that in 2017, more than 900 broadcasting hours in the main Catalan channel were audio described. In the report by Rovira-Esteva & Tor-Carroggio (2018) on Catalan-speaking broadcast, TV3 provided 447 hours of audio subtitled contents (news reports) in 2017.

Spain presented a transitional law (Law on Urgent Measures for the Stimulation of Digital Terrestrial Television on the Liberation of Cable Television and the Encouragement of Pluralism, 2005) which aimed to increase compulsorily the percentages for all the services (divided in subtitling, sign language and AD) in public television until 2015. The law was last modified in 2015 (General Law on Audiovisual Communication, 2015) and percentages are given for AD and subtitling in Spanish broadcast to be reached by 2013 (75% of programs should be subtitled and a minimum of 2 hours a week for audio described programs). The situation remains uncertain since the modification of 2015 does not provide any new information beyond 2013, as it is collected in the Spanish Code of Audiovisual Rights (2017), last modified in November 2017. Regardless of the percentages and minimum amounts of time, neither the Catalan nor the Spanish mention the application of AST.

There exists a tendency towards increasing or applying, in the first place, accessible services for the deaf and the hard-of-hearing in the form of subtitles excluding the specific requirements for the visually impaired. Cyprus highlights the necessity of making news bulletins accessible with subtitles for the deaf and established a transitory law which encourages broadcasters to increase accessible programs by at least 5% (Radio and Television Broadcasters Law, 2016). It also happens in the Flemish community of Belgium (Decreet houdende wijziging van diverse bepalingen van het decreet van 27 maart 2009 betreffende radio-omroep en televisie, 2018), where accessibility is understood only in forms such as subtitles or sign language, even if AD has been broadly studied by one of their scholars, Aline Remael (2005). This reality is found also in countries such as the Netherlands (Media Decision, 2010), even if it is a country where the development of automatic AST has been applied to regular television (Verboom et al., 2002). Greece offers only percentages for subtitles for the deaf and the hard of hearing (Presidential Decree on the Harmonisation with Directive, 2010).

In general, AST is not mentioned directly in all these regulations. Excluding Bulgaria, which only mentions AST as a way of making accessible a TV program (Electronic Communications Act, 2007), the few cases where they are mentioned are Finland and Sweden, as they have the only regulation including the actual implementation of AST service in a governmental document, with even percentages to be accomplished in the next years. The presence of AST in regular TV broadcasting in Nordic countries is reported by the European Broadcasting Union (EBU, 2016), pointing out Denmark, Norway, the Netherlands, Sweden, Finland and Estonia as countries whose broadcasters provide the AST service. Finland, in its Act on Television and Radio Operations (Act on Television and Radio Operations, 2010) aims at providing quota rates for automatic AST programmes of 10-15% in private services and of 50-100% in public broadcasters for Finnish and Swedish language by the end of 2016. Sweden includes a spoken text service to be offered in all programmes that are not live broadcast by 2019 (Culture and Accessibility - Public Radio and Television 2014-2019, 2014).

The line between AD and AST is sometimes very slim. AST may not be directly mentioned but included when discussing reading aloud text on-screen or captions. Since it is difficult to clearly identify these elements it is useful to observe also regulations on AD. This is the main access service used by the blind and visually impaired and its popularity and history has allowed for more presence in media accessibility regulations.

The Czech Republic observes a 2% of programmes with accessible services for blind people in its Act on Czech Government (Act on Czech Television, 1992), and demands also AD for persons with visual impairments in on-demand products (On-demand Audiovisual Media Services Act, 2010) without mentioning AST. The Léotard Law in France (Loi relative à la liberté de communication, 2016) establishes that programmes with more than a 2.5% of share must be made accessible for the blind and the visually impaired. Poland is on an Amendment to the current Act on Radio and Television (Act on Radio and Television, 2018) aims to increment the thresholds for accessibility at 15% of broadcast per quarter in 2019, 25% in 2020-2021, 35% in 2022-2023 and 50% from 2024 a future 50% of the programming content in 2018. Slovakia requests a 20% of programmes with commentary for the blind for public services and a 3% for independent licensed broadcasters in its Broadcasting and Retransmission Law (Law on Broadcasting and Retransmission, 2000). The French community in Belgium has consolidated a new version of its decree on audiovisual media (Decret Coordoné sur les Services Médias Audiovisuels, 2018), in which the 25% of fiction and documentary programs in prime time (13h - 24h) should be audio described in public broadcasters, and a 20% for private broadcasters in programs with more than 2.5% of audience. The percentage is 15% of audio described programs for those emissions with less than 2.5% of audience in public broadcasters.

The UK only enforces accessibility for the three public authorities, BBC, Channel 4 and the Welsh S4C in its Equality Act (Equality Act, 2010). The Code on Television Access Services developed by Ofcom (Ofcom, 2017) provides in general percentages for signing (5%), subtitling (80%) and audio description (10%) for all programs. Licensed public service broadcasters and Welsh S4C and the BBC, are obliged to meet their interim requirements to which they committed. The BBC is required to subtitle 100%, audio describe 10% and sign 5% of their content. Channel 4 must meet a 90% for subtitles, 5% for signing and 10% for AD and S4C , 80%, 5% and 10% correspondingly (Ofcom, 2017).

Portugal, in its turn, states that the public channels RTP1 and RTP2 are obliged by law to broadcast 70 and 20 hours a year of programmes with AD respectively. For the private channels (SIC and TVI) the amount is of 12 hours a year (Entidade Reguladora para a Comunicação Social, 2016); it has to be said, though, that payment-TV providers are not obliged to broadcast AD programmes in Portugal.

Despite the fact that they do not offer exact percentages, the United States, in their Twenty-First Century Communications and Video Accessibility Act of 2010 (Twenty-First Century Communications and Video Accessibility Act, 2010) and the ulterior Implementation of the aforementioned act (Federal Communications Commission of the United States, 2016), provide a list with the top broadcasters and the techniques to be used to make programmes accessible. Due to the administrative organisation of the country, laws change from state to state. However, all the states who introduced codes and regulations on media accessibility have done so by introducing what is said in the Rehabilitation Act (House of Representatives of the United States of America, 1973). For example, the New York state policy on web accessibility reads: “state agencies will provide a video description for multimedia content that contains video […]” (New York State, 2010, Section A.11.2.).

From what has been said so far, the legal status of MA changes from one country to another (see Table 1 in Annex 1). In some cases, higher institutions such as the European Union or the North American Parliament are imposing new and better regulations that ideally should be applied by all countries and states (in the American case). The contents of a new European Accessibility Act is expected to be passed very soon as they have reached a compromise in the Parliament and Council of the European Union. The contents of the soon-to-be passed are considered as directive, thus reaching a legal obligation for all member states (Parliament and Council of the European Union, 2019). As stated by the European Disability Forum (Moledo, 2017), the new act aims to “to ensure that private entities that offer facilities and services which are open or are provided to the public take into account all aspects of accessibility for persons with disabilities” and provide information on the “functions, practices, policies and procedures […] in the operation of the services targeted to address the needs of persons with functional limitations” (n.p.). The new contents of the acts describe audio description and spoken subtitles as audiovisual accessibility services (as stated in the Audiovisual Media Services Directive of 2010). It compels states to pass national laws according to this new directive in the next three years and obliges states to issue reports every 5 years on the application of such directive. States must impose penalties for the unfulfillment of the contents of the directive. Furthermore, a commission will be set for the assessment of the implementation of the directive.

9. Standards and Guidelines: AST and AD

This section aims to provide an overall vision of the contents related to AST in standards and guidelines or recommendations. The main difference between these two types of documents are the emission bodies behind them. A standard is a document issued by an official standardisation agency, either at a national, or international level. Guidelines are recommendations produced by private and public companies, professionals or scholars.

As mentioned previously, the section focuses primarily on an overview of standards and guidelines about AST. However, since AST is usually offered in combination with AD, specific standards and guidelines on AD are also included, putting the focus on what is said about AST and, more generally, about making accessible on-screen text. It has to be said, though, that the majority of such indications are often contradictory (Gerber, 2007) "because of their ambiguity, subjectivity, and also because too much is left to personal interpretation and taste" (Orero & Wharton, 2007, p. 166)

References to AST in standards and guidelines are generally poor and fuzzy. In most cases written subtitles appear in the "What to describe?" section in AD standards and guidelines. The key information provided is that written captions and subtitles must be included in the AD of a film. Some of them are published by national standardisation agencies and regulators, such as in Spain (AENOR, 2005), Ireland (Broadcasting Authority of Ireland, 2012), and the United Kingdom (Ofcom, 2017); and also by international agencies (ISO/IEC 20071-21, 2015). In others cases guidelines are proposed by broadcasters as internal documents, for example the Bayerischer Rundfunk (Benecke & Dosch,

2004) in Germany or in the United States (Netflix, n.d.). Moreover, different associations and institutions distribute their own guidelines. This is the example of Media Access Canada (Milligan & Fels, 2013), the Polish Audiodeskrypcji Fundation (Szymańska & Strzymiński, 2010) and Kultury bez Barier Fundation (Künstler, Butkiewicz, & Więckowski, 2012), the American Council of the Blind (Snyder, 2010) and the Italian Senza Barriere ONLUS association (Busarello & Sordo, 2011). Some scholars have also issued standards and best practices. In Belgium, Remael (2005) and Remael, Reviers, & Vercauteren (2014); in the United Kingdom, O'Hara (n.d.); in France, see Morisset & Gonant (2008); in Spain, for the AD of opera in the Catalan region, Puigdomènech, Matamala and Orero (2008) and finally Georgakopoulou in Greece (2008).

Audio subtitles are included in some of the standards and guidelines for the production of AD. The Spanish Standard (AENOR, 2005) points out that "the script [for AD] should have all the information given by occasional subtitles". The Irish guidelines (Broadcasting Authority of Ireland, 2012) advise that relevant on-screen signs of writing should be described and the same applies for subtitles and captions in current affairs documentaries. Basically, in both rules, subtitles are converted into an oral explanation.

The Guidelines and Best Practices from the American Council of the Blind (Snyder, 2010) recommend audio subtitles to be introduced with "subtitles appear" and a change of intonation in the voice. This information resembles what reads in the Audio Description Background Paper by the Mikul (2010) for Media Access Australia and Georgakopoulou's guidelines in Greece (Georgakopoulou, 2008). The former mentions that AST should be introduced with "subtitle" or "he/she says". In the latter, it is stated the word "subtitle" should be typed before the subtitled text when preparing the AD script, without giving further instruction on the actual voicing.

In the style sheet created by Remael (2005), subtitles are included in the "subtitles or other text on screen (e.g. graffiti), opening titles and/or credits." even if text on screen without a sound input differs from written subtitles that are accompanied by an original audible utterance. The same goes for one of the Canadian guidelines by Accessible Media Inc. in collaboration with the Canadian Association of Broadcasters (2014). In contrast with the previous more elaborated description of the suitable strategy, it only mentions that the nature of the text on screen should be mentioned, i.e. “caption”, “logo” or “subtitles”, among others.

The use of voices is mentioned in the comparative study of guidelines published by the Royal Institute of the Blind People (Rai, Greening, & Leen, 2010). The use of at least two actors, man and woman, is recommended for AST. Ofcom, in its Code of Television Access Services, indicates that “to differentiate between subtitles and description the describer should do this by either the use of their voice (e.g. stating the obvious, ‘He says in Russian…’ or ‘A caption reads…’) or a second voice”. The Canadian guidelines by Media Access Canada (Milligan & Fels, 2013), more extendedly, provide a section only devoted to the voicing of subtitles “when characters speak in a different language from the primary language of the show” (n.p). The way these should be delivered is by “read[ing] the translation after stating that a subtitles appears” (n.p.) and considers the practice as “describing the screen real estate”.

Other documents contemplate a deeper application of AST. In the guidelines issued by the German broadcaster Bayerischer Rundfunck, Benecke & Dosch (2004) point out the recent application of “acoustic subtitles” in films. They advise to introduce text or subtitles, on their first appearance, with expressions such as “words appear” or “subtitles appear” and these to be followed by a usage of a different tone of voice. Remael (2014) in her contribution on audio subtitling to the guidelines provided by the European project ADLAB, provides more prescriptive ideas such as the use of different voices when more than one character feature in the scene or the realisation of different intonations when working with low-budget products. She also presents the two main strategies for voicing AST: dubbing and voice-over effect, being, according to her voice-over the more common strategy. She highlights the fact that for a dubbing effect it is important to make clear character recognition, since the voices from the original dialogues cannot be heard and complete different voices are used. For scenes with little dialogue in a foreign language she recommends issuing a description of the subtitles rather than voicing them. The narration or description of the speech exchange can fulfil the visual information found in the written subtitles.

More recently, the ISO & IEC published its recommendations on the audio delivery of written text and captions on screen (ISO/IEC 20071-25, 2017).

This provides information about both technical and implementation recommendations. It describes the different treatment of volumes (between the original audio and the caption/subtitle delivery), the synchronisation of the elements to be audio-rendered and their corresponding audio excerpt and the prosodic characteristics that the delivery of AST can use to, for example, help character identification. It mentions as well the dubbing and voice-over effect. It is the most accurate document when considering how AST should be presented. Even if not providing strict norms for their implementation, it enumerates the features that intervene and that can be modified to provide the user with different results depending on the characteristics of the product. The ISO standard makes a call for more research on user reception in order to address more accurately some of the characteristics shown. In general, guidelines issued by governmental institutions, associations and broadcasters do not include information on AST and when they do, it is considered part of the AD. Only the ADLAB and ISO & IEC guidelines provide information and this tackles general aspects of the delivery of AST.

10. Conclusion

MA has drawn some attention in the research community within AVT studies. The development of new technologies and supports, added to the increase in social awareness towards the construction of more inclusive environments, has led to more research. Services such as AST have seen more presence in current audiovisual content as well as in the academic field. Either as a part of the AD or as a separate service, AST has proved to be useful for a variety of audiences: primarily the blind and partially sighted audience in and also for audiences with different specific needs related to reading accessibility, from dyslexic to elderly viewers.

AST is presented as an aural rendering of written subtitles. The differences in the creation and delivery of AST allow for a categorisation of audio subtitles. In this chapter, they are classified according to (a) whether they are integrated in and autonomous from the AD track; (b) voiced with synthetic or human voices, and (c) the moment the AST track is mixed: broadcast-mix or receiver-mix. Autonomous AST has been put into practice in some territories and so is reflected in some of the regulations concerning media accessibility. Nevertheless, the inclusion of AST in broadcasting and audiovisual legislation is scarce and normally related to the AD service. Few information is given about the way AST should be made available to the user. On the other hand, guidelines and standards centre their attention on human-voiced AST as a part of the AD track as well. In some cases, these types of documents collect instructions and recommendations on how to deliver AST, examples of this are the ADLAB guidelines (Remael et al., 2014) and the ISO/IEC 20071-25 standard (2017).

When it comes to research, two different paths have been taken. On the one hand, some works have investigated the production of autonomous AST voiced with synthetic voices and for regular TV broadcasting. On the other hand, some studies have centred on human-voiced AST for fiction films. Here, AST has been studied mainly from a descriptive approach. These works present analyses of corpora formed by films already provided with AD and AST services or provide some proposals for the AD and AST of films which contain in their scenes written text on screen (written subtitles). Most of the films in the corpora show more than one language, thus portraying a multilingual reality. In such content, even if the main language spoken in the film is adapted by means of dubbing (in dubbing countries), secondary languages are left with the provision of written subtitles. These subtitles are not accessible for everyone to the same degree and that is the reason AST are observed as the solution for audiences with specific needs.

The portrayal of multilingual realities has increased its presence in current audiovisual production. Various studies have worked on the translation of multilingualism by means of different transfer modes, such as subtitling or dubbing, and also by using access services such as SDH. From this perspective, the treatment of multilingualism in translation proposed in Sternberg's representation of polylingualism (1981), which presents a scale from more to less domesticating approaches in the translation of literature, has been used as a model adapted to AST. The adaptation is based on the recent application of the same representation to SDH. For AST, instead of using visual information to reveal multilingualism, as in the case of SDH, audio subtitles use the aural code to provide a broad variety of forms when homogenising dialogues to different degrees. This is obtained by combining the form of AST and the information provided by the AD track about the multilingualism found in the content. Although different strategies are provided, two main different effects are considered up to date in the literature about the delivery of AST: dubbing and voice-over. The two effects draw their characteristics from these two traditional AVT transfer modes.

Standards, guidelines and research studies have included these two effects in their contents but their differences had not been established in detail yet. A dubbing effect is that in which the original track remains inaudible while the AST is displayed in synchrony and using an acted voice which aims at conveying the prefabricated naturalness established by the practice of this technique in the country audiovisual tradition. Contrastingly, in a voice-over effect, the AST is delivered on top of the original track, which is heard in the background at a lower volume. The AST then is displayed in the isochrony found in voice-over, starting before or after, and with a prosodic pattern resembling that of reading-aloud prosody, i.e. showing less marked prosodic features.

This theoretical framework serves as a general basis for the next sections in this thesis, focused on user reception of AST. It presents a general landscape of the AST service in both regulations, and standards and guidelines, and in AVT studies research. Furthermore, it collects the definitions of AST and their categorisation, as well as the description of their features and the forms they can have. In fact, the information provided in the AST can influence the treatment of multilingualism, key matter in this research work. The description of the two different effects, dubbing and voice-over, is of crucial importance, since they are contrasted in our experiments.

Chapter 3

User Reception:

Theories and Methods

Chapter 3
User Reception: Theories and Methods

The objective in this Chapter is to research emotions in media environments, which are viewed as a source of entertainment. In this regard, various concepts are used in the literature in relation to user experience. The main approaches to these concepts are reviewed in order to build up both a theoretical framework for the concept of emotion and a methodological framework for the experimental part in the thesis (see Chapters 6 an 7).

Section 1 provides a general overview of the study of the reception of media within the discipline of AVT and MA. **Section 2** describes the process of media consumption as entertainment, and will relate it to the theories of flow, presence and, ultimately, to emotion. Apart from mentioning to previous research on media studies, reference to specific studies in AVT and MA are also made. **Section 3** links of the theories to the measurement of emotional activation. **Section 4** defines what an emotion is and provides a categorization of emotions, proposing a hybrid approach for the framework of this thesis. The type of emotion used in the experiments is described and analysed in more detail. **Section 5** deals with methodological aspects and describes the various tools that have been used to measure emotions. The emphasis is put on studies in which audiovisual stimuli have been used, and reference to specific MA research is highlighted (**Section 6**). The chapter will conclude with some final remarks on the theories observed in this framework.

1. User Reception in Audiovisual Translation and Media Accessibility

Extensive research on the reception of the main transfer modes in AVT, dubbing, subtitling, and voice-over, has been carried out in recent decades. To name just a few, see Matamala, Perego, & Bottiroli (2017) and Perego, Del Missier, & Bottiroli (2014) for dubbing and subtitling; Perego, Del Missier, Porta, & Mosconi (2010), Kruger et al. (2016) for subtitling, and Sepielak (2016) for subtitling and voice-over. Particularly relevant are the works regarding AD, closely linked to the thesis' object of study, where reception research is also quite extensive (see, for instance, Fernández, Martínez, & Núñez, 2015).

The reception of the AD service has been studied when applied to theatre (Udo et al., 2010), fashion shows (Udo & Fels, 2009) and opera (see Cabeza i Cáceres, 2010; Eardley-Weaver, 2014; Matamala & Orero, 2007; Orero, 2007a; Orero & Matamala, 2007). In the AD of cinema many scholars have produced interesting research. For instance, audience preferences were studied when it comes to the information provided by the AD in the form of AI before the commencement of the film (Di Giovanni, 2014a; Jankowska, 2013; Romero-Fresco & Fryer, 2014b) and it was also applied to the reception of audio described Bollywood films (Rai et al., 2010). The amount of information processed and retained by blind audiences by means of AD was studied in Fresno Cañada's thesis (2014). Taking a different approach, the way AD is delivered was also tested with final users in terms of speech rate and its impact on user's comprehension (Cabeza-Cáceres, 2013) or the characteristics of synthetic and human voices and consequent impact on the user (Fernández-Torné, 2016; Fernández-Torné & Matamala, 2016; Szarkowska, 2011; Szarkowska & Jankowska, 2012; Szarkowska & Mączyńska, 2011; Szarkowska & Walczak, 2012; Walczak, 2017; Walczak & Fryer, 2018). Particularly relevant to the focus of this thesis is the study carried out by Thrane (2013) on the elaboration and reception of AST delivered with synthetic voices in Sweden.

The present work intends to take a step further following the line of studies on user experience by applying new instruments targeted at more objectivity. This new methodology is based on theories on mental processes and the responses they provoke in the organism. Paying special attention to the branch of MA, studies with visually impaired participants and filmic stimuli are found in the notion of presence (Fryer, 2013; Fryer & Freeman, 2014; Walczak, 2017; Walczak & Fryer, 2018; Wilken & Kruger, 2016) and that of emotional activation (Ramos, 2015; Ramos, 2016). Such concepts are explained in following sections specifically applied to the media and audiovisual content.

2. Media as Entertainment

The reasons why audiences use media and the expectations derived from their use have been broadly discussed by scholars (see Sherry, 2004). In an attempt to shape explanations behind the captivation created by media, different concepts arise, such as flow, enjoyment, presence or emotion. These concepts are in some cases not as clearly defined and categorised as the scientific community may wish. For instance, Lyons et al. (2014) point out that for many authors terms such as presence and engagement are interchangeable while there is a difference. Similarly, Kreibig (2010) stresses the lax and inconsistent use of terminology in the field of emotions, and Jsselsteijn, de Ridder, Freeman, & Avons (2000) suggest that the terminology for the factors determining presence “tend to vary across authors” (p. 521). The concepts are narrowed down in the next sections within the framework of the study.

Ruggiero (2000) argues that one of the first uses of media is diversion (understood as entertainment), which is exemplified by the author as “an escape from routines or for emotional release” (p. 26). The idea of audiovisual media as a way of escapism is also exploited in Vorderer & Hartmann (2008) quoting early scholars working on the topic (Katz & Foulkes, 1962) when describing the beginnings of the study of media entertainment, now deemed to be an established field in communication science and media psychology. Entertainment is understood as a “positive moodlike meta-emotion, which arises from the appreciation of underlying primary emotions” (Vorderer & Hartmann, 2008, p. 533). Entertainment happens in a context where the user is located in an environment that should be understood both outside and inside the audience. Thus, it is approached from an ecological perspective where the user exists in a real world but is involved in a mediated environment (Bryant & Miron, 2002).

According to Vorderer & Hartmann (2008), people employ media to regulate their excitation. By means of reappraisals, users can balance primary emotions and achieve pleasure, even if such emotions are not pleasurable (see Carroll’s Philosophy of Horror, 1990), becoming the central element of the phenomenological experience of cinema[4]. As a matter of fact, “pleasure [is] the oldest and still key function regulating human behaviour” (Miron, 2006, p. 344), it is “the key to feeling entertained” (Vorderer & Hartmann, 2008, p. 535). Plantinga (2009) and Vorderer & Hartmann (2008) suggest that the explanation behind the entertainment in media consumption is found in the theories of

[4] This leads to the enjoyment of emotions such as fear (see Cantor, 2002), pain, and sadness (Rozin, 1999; Rozin et al., 2013; Schramm & Wirth, 2010).

media flow and presence. The audience experiences “emotions or quasi-emotions within the context of a game or make-believe” (Plantinga, 2009, p. 88). Those exposed to an audiovisual content by which they are entertained, enter in what Vorderer & Hartmann (2008) describe as “involved mode”. That means that the user starts to “think ‘within’ the depicted mediated world” (p. 536).

In the literature the terms “entertainment” and “enjoyment” can be used as synonyms. In this chapter, the term used is that of entertainment. The explanation is found in the thin line delimiting what is and is not enjoyable or pleasurable. On this topic, Carroll (1990) discusses on the enjoyment of topics loaded with negative emotions such as horror, which results in some audiences’ enjoyment. Entertainment, thus, seems to cover a larger variety of experiences which are based in the emotions transmitted.

This being said, for entertainment to happen, the spectator should feel immersed in the content which provides emotions that will ultimately bring them to a pleasurable state. The terminology used to describe this immersion varies across theoretical models but the focus divided into two concepts: the concept of flow, as described in Csikszentmihalyi (1988) and applied specifically to media by Sherry (2004), and the concept of presence, as described by Lombard & Ditton (1997). The relationship between these two concepts and the concept of emotion is discussed in the next sections.

2.1. Media Entertainment: Flow

Sherry (2004) disentangles the notion of flow, first developed by Csikszentmihalyi during the 70s (see Csikszentmihaly, 1988; Csikszentmihalyi, 1988, 1993, 1997) and applies it to media, particularly to film and television. Flow is primarily related to the pleasure experienced from the immersion in different activities and tasks. He highlights the entertaining role of media and points it as one of the most important gratifications sought in media. To him the interest does not reside in the fact that people are entertained by media. The discussion arises when considering the way people are entertained.

Before the theory of flow was applied to media, the theory developed by Csikszentmihalyi (1997) was created as universal and applied to different tasks. He described flow as a self-motivating state which was characterised by a list of features. Among the features described, some can be related to the experience generated by “traditional” audiovisual products (i.e. cinema and television): an intense concentration in what is being done, the loss of self-consciousness, the distortion of time, and the perception of the experience as rewarding (see Nakamura & Csikszentmihalyi, 2002). Besides, for a stimulus to provoke the experience of flow, there should be a clear balance between an individual’s skills and the difficulty of the activity one faces. In other words, failing to understand or experience in full an activity will bring to a decay in entertainment. Even if not directly applied to sensory disabilities, Sherry (2004) points out the impossibility to understand a language as one of the factors that hinders immersion and entertainment. This highlights the importance of this type of research in the fields of AVT and MA.

Both the definition of flow provided in Csikszentmihalyi’s works and the later application to media by Sherry concern a particular feeling of evasion. This feeling, linked to the features of the flow described in the last paragraph, finds comprehensive explanation in the phenomena of presence and transportation (Lombard & Ditton, 1997).

2.2. Media Entertainment: Presence

The link between the theory of flow and the experience of the phenomenon of presence has been commented by some of the main scholars in the field (Lombard & Jones, 2015). As a broad description, presence can be explained as “an illusion that a mediated experience is not mediated" (Lombard & Ditton, 1997). In some way, the concept of presence when applied to audiovisual media refers to the immersion produced by a non-real and mediated input and the experience of transportation inside the mediated environment without being physically moved.

The concept of presence is still being debated and perfected, particularly in the terminology used and the items that can or cannot be considered factors causing an impact to the feeling of presence (Lessiter, Freeman, Keogh, & Davidoff, 2001). The lack of standardisation of the term arises from the different conceptions of presence across the scientific community. As a matter of fact, some of the main scholars specialised in the topic are calling for for the terminology to be expressed as precisely and consistently as possible as a way of scientific establishment (Lombard & Jones, 2015).

Lombard & Jones (2015) provide a categorisation of the different aspects involved in the notion of presence studied and identified by different scholars. One of the first aspects of presence that have become one of the key aspects studied from the incorporation of technologies such as virtual reality (VR) is spatial presence, that is, the responses to spatial cues and the mental models of mediated spaces that create the illusion of place (Biocca, Harms, & Burgoon, 2003, p. 459). Spatial presence has always been related to the term telepresence or transportation, referring to the arrival and return from a mediated place or space (Biocca & Kim, 1997). Social presence is another aspect. It could be defined using the words in Reeves & Nass (1996, p. 5) as presence experienced in the "individuals' interactions with computers, television and new media […] just like interactions in real life". Another aspect is related to the perception of the self: self-presence. When applied to media, it refers to "the extent to which some aspect of a person's proto (body schema) self, core (emotion-driven) self, and/or extended (identity-relevant) self is relevant" (Ratan, 2013, p. 336).

Other aspects in the categorisation by Lombard & Jones (2015) that are very relevant to the interest of this research are engagement and realism. According to Freeman (2004), one of the key dimensions of presence is what he refers to as "engagement/involvement/attention" (n.p.). Lombard & Ditton (1997) point out in other words, that "when the users feel immersive presence they are involved, absorbed, engaged, engrossed" which bridges with the characteristics of media flow as described by Sherry (2004). Even if the terminology may be diverse, it can be argued that the idea is shared by many scholars. Realism is also considered, according to Lee (2004, p. 34), a "psychological state in which virtual (para-authentic or artificial) objects are experienced as actual objects". Freeman (2004) situates realism as the second of the three dimensions of presence, described as "naturalness /realness/consistency with the real world". This dimension is also referred to as "ecological validity" in Freeman, Lessiter, Pug, & Keogh (2005) based on other works by Freeman (2003, 2004). Realism is also portrayed in Csikszentmihalyi's flow theory (1997) when he argues that the difficulty in interpreting messages that fall far from the formal norms is linked to the lack of the experience of flow. The mediated message, as Sherry (2004) puts it, which does not follow the formal characteristics, will be less accessible and create obstacles in the process of entertaining and creating enjoyment.

3. Entertainment, Flow, Presence and Emotional Activation

The connection of the concepts presented in this theoretical framework (entertainment, flow and presence) has been mentioned in previous studies. Weibel, Wissmath, Habegger, Steiner, & Groner (2008) carried out a study on virtual reality in which participants were asked to play online games both against machine-controlled and human-controlled opponents. In the study, data were collected by means of self-report instruments and provided evidence suggesting that the three processes were strongly related. They are also related to emotion in Dillon (2006), who actually highlights the fact that research on the link between presence and emotional activation has not been well researched. Furthermore, she insists on the need for more research on the connection between presence and subjective and psychophysiological measurements.

The direct relation between the level of emotion and that of the experience of presence has been discussed by other scholars. Riva (2011) simply argues that the higher the level of presence the higher the emotions the user is experiencing. The link between presence and emotion, though, is according to some scholars only applicable when the user is exposed to immersive virtual reality environments, and more specifically, to the treatment of phobias, fear being one of the strongest inducers of arousal (Diemer et al., 2015). In the study by Meehan et al. (2002), known as the Pit Room, users with phobia to heights were immersed in a virtual environment which recreated that feeling. Furthermore, these studies used psychophysiological measures, thus implying a fundamental link between emotion and changes in the organism. Some studies base their studies on the idea of presence as a phenomenon triggering physiological changes[5] (see Section 5.2).

[5] See, for example, Baños, Botella, Alcañiz, & Liaño, 2004; Dillon, 2006b; Dillon, Keogh, Freeman, & Davidoff, 2001; Freeman, Avons, Meddis, Pearson, & IJsselsteijn, 2000; Freeman, Lessiter, Pugh, & Keogh, 2005b; Fryer, 2013; Fryer & Freeman, 2014; IJsselsteijn, de Ridder, Freeman, & Avons, 2000; S. Lee, Kim, Rizzo, & Park, 2004; Meehan, Insko, Whitton, & Brooks, 2002; Mikropoulos, Tzimas, & Dimos, 2004; Regenbrecht, Schubert, & Friedmann, 1998; Usoh et al., 1999

Whether or not less immersive environments could lead to the same type of reactions was also a challenge in academia (Meehan et al., 2002). In this regard, Dillon et al. (2001) run experiments to test the effect of monoscopic and stereoscopic videos on users in terms of presence. They took the inspiration for their research on previous studies in which the correlation of measures of psychophysiology was discussed (see Detenber et al., 1998; Reeves et al., 2009; Simons et al., 1999). These studies included the use of motion or still pictures in relation to emotional activation and attention.

Moreover, the sense of presence is also increased by the affective impact of the content, translating into changes in the physiological characteristics (Freeman et al., 2000; IJsselsteijn et al., 2000). Video clips which were able to elicit emotional activation proved to provide higher levels of presence than other materials considered to be neutral (Dillon, 2006). This parallelism is also discussed in Baños et al. (2004), who made use of virtual environments in three different modes showing two emotional conditions (sad and neutral). In this study, the three different modes of presentation were compared (PC monitor, a projection in a wall, and head-mounted display) with emotional and non-emotional content in terms of immersion and affective content through self-report instruments. They conclude that both immersion and affective content have an impact on presence. In a more theoretical approach, the interrelation between presence and emotional content in media is discussed in Freeman et al. (2005). They highlight that presence and emotion will be somehow connected only if the contents are arousing, that is, if they are relevant to the consumer of the content. They point out that features such as screen size or navigation in virtual environments (VE) have a direct impact on the feeling of presence. Nevertheless, according to their experiments, they did not find changes in the participant’s emotional responses. Yet, they highlight the complexity of the phenomenon and advocate for the coexistence with the idea that presence is linked to affectivity of the stimulus.

Directly connected with the experiments carried out in this project, some studies with blind and partially sighted participants, researchers also relied on the emotion in the scenes selected for the presentation of stimuli. Ramos (2015) made use of AD clips portraying the emotions of sadness, disgust and fear. Fryer (2013), though focusing on the study of presence, linked emotion as one of the factors provoking physiological changes, which she measured to complement the data extracted from self-report tools.

4. Emotions

This theoretical framework sustains that the concepts here explained such as entertainment (Bartsch, Mangold, Viehoff, & Vorderer, 2006; Klimmt & Vorderer, 2003) and flow (Michailidis, Balaguer-Ballester, & He, 2018; Sherry, 2004) or presence (Dillon, 2006; Freeman et al., 2005; Fryer & Freeman, 2014) related to emotion. Research on emotion has increased in academia in last decades (Rottenberg et al., 2007). The study of emotions is referred to with different terms, which also reflect different approaches: terms such as emotionology (Stearns & Stearns, 1985), emotioncy (Pishghadam, Firooziyan, & Esfahani, 2017), affective neuroscience (Lindquist, Kober, & Barrett, 2015) or affective science (Rottenberg, 2003) have been introduced in different disciplines in recent years (Izard, 2010; Serrano-Puche, 2016; Ticineto Clough & Halley, 2007). Likewise, new methodologies are proposed for their measurement, such as the study of psychophysiological responses, which seems to be a flourishing practice in many fields (Kreibig, 2010). As stated by Mauss & Robinson (2009), there is not a "gold standard" for their measurement. However, it is true that progress in the field is being made and that some theories are commonly shared by the community.

First of all, this section aims to define and categorise the concept of emotion. Secondly, two different models of categorisation of emotions are presented, namely the discrete and dimensional approaches to emotions. It is worth anticipating that both theories can find reconciliation, and that the theoretical approach here consists of a combination of both. Thirdly, the elicitation of certain particular emotions are analysed considering the purposes of Study 2. An explanation concerning the relation between physiological measures, particularly those considered here, namely electrodermal activity (EDA) and heart rate (HR), and how these gather physiological information are also provided.

4.1. Defining Emotion

A distinction must be made between emotion and other concepts normally related to the concept in regular conversation such as feeling, affect or mood. Feeling is defined as “a single component denoting the subjective experience process” of emotion (Scherer, 2005, p. 699); in turn, the affect or mood is understood as a continuous, longer and slower moving state (Watson, 2000), or as “a diffuse and long-lasting affective state that is not object-related, i.e., not experienced in simultaneous awareness of its causes” (Kreibig, 2010, p. 410), see Frijda, 1993; Gendolla, 2000).

In opposition, emotions are conceptualised as “bursts of activity, or waves, superimposed against the backdrop of his affective stream” (Rottenberg et al., 2007: 10). This distinction, though, helps solely to grasp some information about the duration and the volubility of an emotion. However, defining the concept of emotion proves extremely complex. To begin with, and as stated by Widen & Russell (2010), an emotion can be observed from two different approaches which, in turn, develop two different concepts. The term emotion in regular language and daily life is qualified by Izard (2010) as a descriptive definition in opposition to the term emotion used in science for research purposes, referred to as a prescriptive definition.

Many definitions have been proposed for the concept of emotion during the last century. At the beginning of the 80s, the exhaustive analysis by Kleinginna & Kleinginna (1981) extracted and categorised 102 definitions of emotion from experts in the field. As many as 11 groups were created: affective definitions, cognitive definitions, those based on external emotional stimuli, those related to the psychological mechanism of emotions and those concerning expressive behaviour, disruptive definitions, adaptive definitions, multi-aspect definitions, restrictive definitions, motivational definitions and sceptical statements (questioning or denying the concept itself). Such is the variety that some scholars demand to stop using the concept of emotion without contextualizing it, and providing a statement of “the meaning or meanings assumed by the author” (Izard, 2010, p. 368). For the purposes of our research, a definition of emotion has been based on previous definitions presented by the following scholars.

Plutchik (2009), describes an emotion as an "inferred complex sequence of reactions to a stimulus, and [that] includes cognitive evaluations, subjective changes, autonomic and neural arousal, impulses to action, and behaviour designed to have an effect upon the stimulus that initiated the complex sequence" (p. 217). Likewise, according to Scherer (1987, 2001, 2005), an emotion is an "episode of interrelated, synchronised changes in the state of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism". Finally, and more related to our film elicitation of emotion, following Lang's proposal (1978), an emotion is described in Rottenberg et al. (2007, p. 10) as

> a transient but coordinated set of responses that occur when an individual faces a situation (real or imagined) that is relevant to salient personal goals. […] We view emotions as being multi-componential, typically involving changes in cognitive, experiential, central physiological, peripheral physiological, and behavioural response systems.

It seems that, despite superficial differences, each definition of emotion characterises it as a series of reactions or responses resulting from a series of stimuli. Izard highlights the brain and nervous system as both recipient of stimuli and sender of responses (2010, p. 367):

> Emotion consists of neural circuits (that are at least partially dedicated), response systems, and a feeling state/process that motivates and organises cognition and action. Emotion also provides information to the person experiencing it, and may include antecedent cognitive appraisals and ongoing cognition including an interpretation of its feeling state, expressions or social-communicative signals, and may motivate approach or avoidant behavio[u]r, exercise control/regulation of responses, and be social or relational in nature.

Finally, and holding consistency with the previous definitions, Kreibig (2010), in her analysis of up to 134 publications on experimental investigations, defines an emotion as a "multi-component response to an emotionally potent antecedent event, causing changes in subjective feeling quality, expressive behaviour, and physiological activation” (p. 410).

Despite the differences in existing definitions, many of the previous authors share common notions that bring the definition closer to a prescriptive rather than a descriptive approach. According to the contents in these paragraphs, and inspired by the previous definitions (Izard, 2010; Kreibig, 2010; Plutchik, 2009; Rottenberg et al., 2007; Scherer, 2005), an emotion is understood here as the response of the organism, both at the behavioural, social and neurophysiological level, provoked by a single stimulus or a series of stimuli that can be both internal or external. The stimuli are processed by our cognitive systems, which are responsible for distributing the information that generates a variety of responses.

After defining the concept of emotion, it is important to understand the way emotions are represented in relation to others. This allows for a better understanding of their differences and similitudes and the way these are cognitively processed and influence the organism.

4.2. Categorisation of Emotions: Discrete and Dimensional Approaches

The categorisation of emotions is defined, first of all, by the prescriptive and the descriptive approaches to emotions (Widen & Russell, 2010). The descriptive approach understands emotions based on the definition normal language would use, that is, how in daily life emotion is characterised. On the other hand, a prescriptive approach will try to define emotions and their characteristics for their use in science, both at a theoretical and at an empirical level (psychologically and physiologically, for instance).

The scientific study of emotions, as discussed before, consists of more than a century of practice with works published as early as 1884 (James, 1884). Similar to the way emotions have been understood and defined during history (Kleinginna & Kleinginna, 1981), different ways of categorising them have been proposed. Such debate brought scholars to take a stance in the dichotomy between discrete and dimensional approaches. In contemporary times, these two approaches are the most considered as recent literature shows (see, for instance, Christie & Friedman, 2004; Fontaine, Scherer, Roesch, & Ellsworth, 2007; Harmon-Jones, Harmon-Jones, & Summerell, 2017; Lang, 1988; Mauss

& Robinson, 2009; Sokolov & Boucsein, 2000; Stadthagen-González, Ferré, Pérez-Sánchez, Imbault, & Hinojosa, 2017). It is worth remarking that both approaches are not exclusive and that a combination of both can be also adopted as defended by Harmon-Jones et al. (2017).

4.1.1. Discrete Approach

A discrete approach to emotions considers that there exists a variety of emotions that are universally understood and accepted. Some scholars highlight the fact that each emotion "refers to a 'family' of states that share distinctive universal signal, psychology, antecedent events, subjective experience, and accompanying thoughts and memories" (Lindquist et al., 2015, p. 2). This approach sustains that there is a finite number of basic emotions. Depending on the author, emotions can be grouped under a different number of basic emotions. Ekman (1992, 2005), for example, distinguishes six different basic emotions: happiness, sadness, fear, surprise, disgust, anger; whereas Plutchik (2001) presents four pairs of basic emotions based on contraries in his theory: joy-sadness, anger-fear, trust-distrust, surprise-anticipation. The rest of emotions are just part of these emotions that have specific and intrinsic characteristics that distinguish one from the other. As put by Plutchik (1980, p. 6), "emotions are rarely if ever experienced in a pure state. Typically, any given situation created mixed emotions which are difficult to describe in any simple or unequivocal way".

This approach has been adopted in recent studies such as in Stadthagen-González et al. (2017), in which they reduce emotions to 5 basic categories when identifying Spanish words to refer to emotions. Schaefer, Nils, Philippot & Sanchez (2010) also consider six emotional discreteness scores after validating a database of films for emotion elicitation. However, they also consider 15 "mixed feelings" that may fall between discrete emotions. Lindquist et al. (2015), in the field of neurology, try to identify specific regions of the brain activated when exposed to certain discrete emotions. It can be argued that this approach is still considered and followed in a variety of scientific disciplines.

By taking a Darwinist approach, such discrete approach is based on the fact that each emotion causes a certain response that relates to a certain evolutionary need, in the sense that "a specific emotion feeling recruits particular types of cognitive and motoric correlates" (Ackerman, Abe, & Izard, 1998, p. 86). As a matter of fact, such identification of common or prototypical emotional patterns is not only seen across humanity but across other species (Plutchik, 1982). Discrete emotions are therefore independent in the sense that they do not depend on the cognitive process of the organism experiencing it but they automatically influence consciousness (Ackerman et al., 1998).

Nevertheless, this model can be deemed too narrow for certain applications. Kim & André (2008, p. 2071) suggest that "this method is that the stimuli may contain blended emotions that cannot be adequately expressed in words since the choice of words may be too restrictive and culturally dependent". For certain studies, such in word validation of emotions (Stadthagen-González et al., 2017), such narrowness or rigid classification offered by this approach can be very helpful. The present study, on the contrary, relies, in addition to this discrete approach, on a wider approach, that presents different dimensions and a more colourful spectrum of emotions.

4.1.2. Dimensional Approach

Several types of dimensional approach have been proposed over the years. Most accurately, such approaches share the concept of dimension but show different types or levels of dimensions. The dimensional approach argues that emotions can be placed along some basic dimensions or axes that organise them. "The two most commonly cited dimensions are arousal and hedonic valence, and the third, less frequently used dimension, is dominance" (Detenber et al., 1998, p. 115). As a matter of fact, Lang (see, for instance, Bradley & Lang, 1994; Lang, 1988) basing his theory on the 19th century German psychologist, Wundt (1896), who labelled three different emotions: *lust* (pleasure), *spannung* (tension), and *beruhigung* (inhibition), introduces the three dimensions previously described: arousal, valence and dominance.

The valence dimension puts in contrast the states of pleasure and that of displeasure or disgust; the second dimension, arousal, presents the extremes of low arousal and high arousal (Mauss & Robinson, 2009, p. 2). The third dimension proposed by Lang (Bradley & Lang, 1994), that of dominance, is described as the "degree of control exerted by a stimulus" (Stadthagen-González et al., 2017, p. 1), in other words, it places as contraries the extremes of being dominated by the emotion or dominating the emotion. However, this third emotion tends to be less frequently included in modern models of emotion categorisation (Lang, Bradley, & Cuthbert, 2008; Montefinese, Ambrosini, Fairfield, & Mammarella, 2014).

Another dimensional concept is that of approach-avoidance, which is postulated by some scholars (Harmon-Jones, 2003; Harmon-Jones, Harmon-Jones, & Price, 2013). Basically, this dimension draws a parallelism between a positive effect, or “appetitive behaviour” (Adams, Ambady, Macrae, & Kleck, 2006, p. 180) of an emotion with the innate conduct of approaching it, which contrasts with the feeling of avoidance provoked by a negative effect, or “aversive behaviour” (p. 108), which is based on the suggestion of Schlosberg (1954), back in the 50s, that a 3D model was required in order to show the “action tendency and the action disposition” (Kim & André, 2008, p. 2070). The interrelation of dimensions is easily identifiable, though. This behavioural dimension of emotions is tightly related to the dimensions of valence and arousal described by Lang (1995). As a matter of fact, he bases his categorisation of dimensions in accordance with the behavioural systems that define approach or withdrawal. In that sense, valence will account for the directionality of the behaviour and arousal for the variations in the activation of such behaviours.

Although it is true that some scholars provide “robust evidence that more than two dimensions are needed” (Fontaine et al., 2007, p. 1054), the application of one model or another depends profoundly on the type of study and the methodology to be used. For instance, Fontaine and colleagues, who were doing research on semantics and the use of emotionally-loaded words for the definition and distinction of emotions in three languages, present four dimensions. However, and for the present study, which aims to quantify physiological data, two main dimensions that are common in all the models here presented are going to be considered: valence and arousal. These two dimensions are frequently incorporated in studies showing two, three or more emotions. Moreover, they are common to all the models whether or not they include the dominance and/or the approach-avoidance dimensions.

4.1.3. Combining both Approaches

Researchers have been swinging between the two main interpretations of emotions: one in which emotions are placed along dimensions and the other establishing each emotion as a discrete entity (Harmon-Jones et al., 2017). However, many studies have taken a half-way direction in the debate by using part of both approaches. Of special interest is the recent work by Harmon-Jones et al. (2017), who contributed with an analysis of previous research where both approaches are considered. According to these authors, both discrete and dimensional concepts are needed to explain the results in research since within the discrete approach emotions are predicted depending on their classification on a positive-negative scale or dimension (p. 4).

The combination of the two approaches was used in studies where filmic stimuli were used. In Schaefer et al. (2010), a database of emotional scenes is validated by means of two self-report instruments: the PANAS (Watson, Clark, & Tellegen, 1988), based on a dimensional approach, and the Differential Emotions Scale (Izard, Dougherty, Bloxom, & Kotsch, 1974), based on a discrete approach. Likewise, in Fernández Megías, Pascual Mateos, Soler Ribaudi & Fernández-Abascal (2011), self-report instruments for both dimensional approach (Self-Assessment Manikin, SAM) and discrete approach (Discrete Emotions Questionnaire) were used to perform a validation of films dubbed into Spanish extracted from previous studies were the same scenes were validated in the original English (Gross & Levenson, 1995; Schaefer et al., 2010).

Contrarily, whereas some studies in which psychophysiological measures were used with emotional stimuli have adopted a pure dimensional approach, others provide less importance to the way emotions are delimited. Examples of the former are found in Detenber et al. (1998), who studied emotion elicitation through motion images with self-report questionnaires, and psychophysiological measures such as heart rate and skin conductance. Psychophysiological measures found their correlation with the arousal dimension in the self-report instrument. The general results showed that motion images enhanced the experience of the participant, that is, positive images were perceived as more positive and the same happened with negative ones. In their review on the current state of the field of psychophysiology of emotions in children, Sohn, Sokhadze & Watanuki (2001) defend the use of a two dimension model for the mapping of emotions (valence and arousal) as it has proved to be effective for the understanding of the engagement of the autonomic nervous system, provoking psychophysiological responses, in emotion (p. 55).

On the other hand, other psychophysiological investigators attached less importance to the way emotions were viewed. This way, there is no difference between taking a discrete approach or a dimensional approach to emotions. The interest lies in the specification of the characteristics of each emotion or dimension by providing a pattern for the physiological changes associated with each of the emotional variables (Simons et al., 1999). However, such assumption, though partly true, is not the most observed in the field. Previous experiences in psychophysiology have considered a hybrid of the two approaches, since it is “the most fruitful approach for ANS specificity research” (Levenson, 1988), taken more recently in Christie & Friedman (2004).

By inducing emotions through musical stimuli, Kim & André (2008) based their study on a mixed-model approach. They used four different psychophysiological measurements: electromyogram, electrocardiogram, skin conductivity, and respiration changes. They aimed at recognising emotions through physiological changes in comparison to other channels such as facial expression or speech. Eventually, this hybridity of approaches allows for a categorisation of certain responses in different emotional states. They devised a four-dimension model which, according to them, provides greater accuracy.

Based on the structure and the aim of the experiments here proposed, this project is going to follow a hybrid approach that will draw on concepts from both discrete and dimensional approaches:

1. The first phase in Study 2 of the experimentation process begins with an emotional validation of clips. We aimed at discerning discrete emotions that had to portray a delimited array of characteristics understood “universally”. In this phase, the presentation of a scene showing a targeted discrete emotion was combined with a two-dimension questionnaire presenting valence and arousal (see, Chapter 6).

2. The second phase takes a dimensional approach. It uses a two-dimension questionnaire in combination with psychophysiological measures based on the dimension of arousal (low and high). However, it can be argued that the approach is not solely dimensional, since the stimuli selected were identified by understanding emotions as discrete entities (see Chapter 6).

Figure 1 expresses in a diagram the concepts that have been presented in this chapter and their relations. The diagram describes how the concept of media as entertainment is linked to the ideas of flow, presence and emotions. The latter is then divided in the two approaches here suggested, which are finally linked in the hybrid approach taken in the development of the experimental part of the thesis.

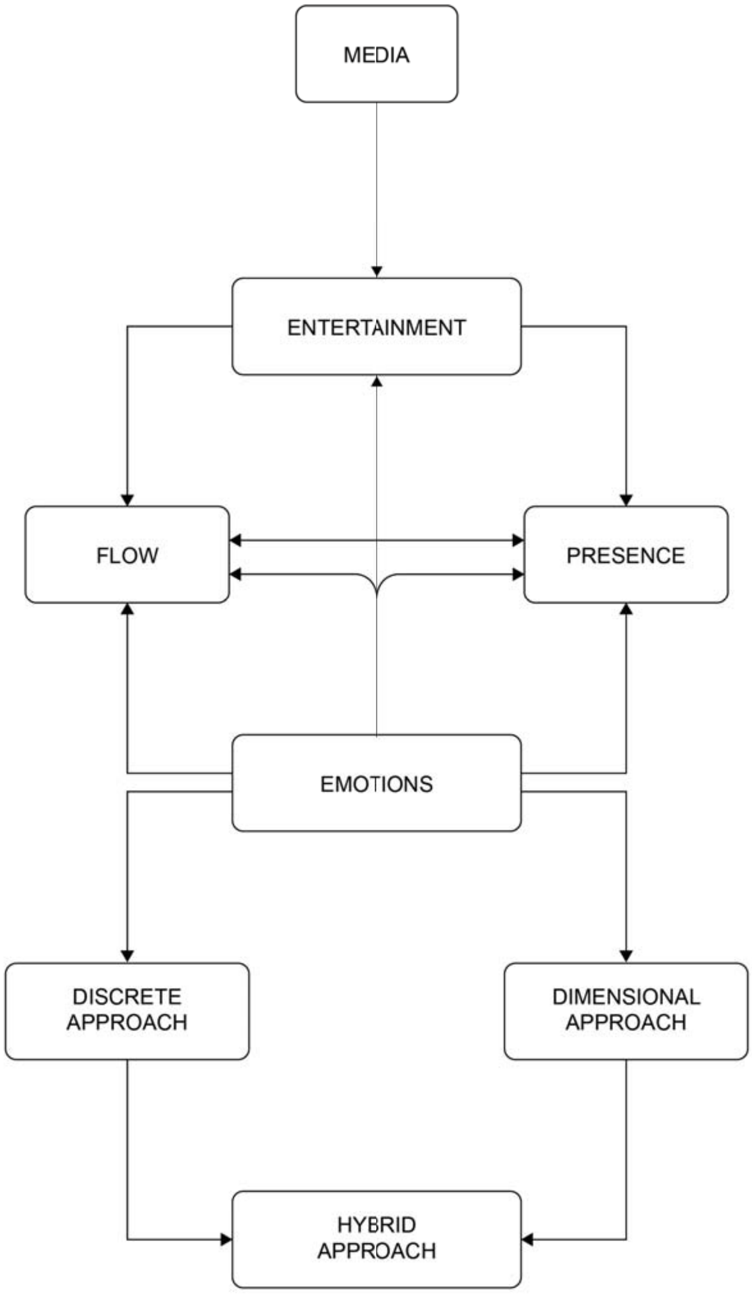


Figure 1. From entertainment to emotion

4.2. Emotions Put to the Test

This section elucidates the placement of discrete emotions in the spectrum considered by the dimensional approach and the psychophysiological responses provoked by specific discrete emotions. It is worth highlighting first the importance of the cultural and semantic understanding of emotions (Montefinese et al., 2014; Stadthagen-González et al., 2017). Undoubtedly, nuances exist among languages when referring to emotional states. However, and as stated above, the following classifications are read as emotional "families", where the label works as a hypernym embracing the many differences that all hyponyms may pose. The universality attributed to such basic emotions, as described for instance by Ekman & Cordaro (2011), provides cohesion and is a required step for the labelling of emotions according to the purposes of the research. In the present thesis the emotions considered are anger, fear and sadness (Silverman & Fels, 2002; Simons et al., 1999).

Such emotions are used based on the psychological and physiological responses they might induce. Some scholars claim that negative emotions or affects are easier to induce in comparison to positive ones and create stronger reactions in the organism (Taylor, 1991; Uhrig et al., 2016; Westermann, Spies, Stahl, & Hesse, 1996). The reason why positive emotions such as joy or happiness are omitted in this thesis is due to in the fact that they have been proven more difficult to delimit. Some researchers claim that this can lead to misconceptions since "a relatively wide range of positive emotions is commonly subsumed under the umbrella term 'happiness'" (Kreibig, 2010, p. 407). The results on a pilot study within the framework of this thesis did not show logical and evidenced results (for more insight in the pilot study, see Chapter 6).

The emotions of anger, fear and sadness will be detailed in the next paragraphs in terms of dimensional characteristic and physiological changes, A table (Table 1) summarises the main features described. Furthermore, the concept of neutral when applied emotion will be provided. Even if not considered one of the basic emotion, if we were to imagine emotions in a spectrum of colours, neutral had to be imagined as black, in opposition to white, that is, the lack of colour. The inclusion of neutral has been suggested in previous studies on emotion with the use of psychophysiological measures (Kreibig, Wilhelm, Roth, & Gross, 2007).

4.2.1. Anger

Anger is considered a basic emotion by all scholars that have issued a discrete categorisation of emotions (see, for instance, Ekman, 2005; Plutchik, 2009). Other emotions exist around and within anger (Harmon-Jones et al. 2017): according to Goleman (2006), anger is related to the following secondary emotions: fury, outrage, resentment, wrath, exasperation, indignation, vexation, acrimony, annoyance, irritability, hostility, and, perhaps at the extreme, pathological hatred and violence. Mentioning the secondary emotions related to the main emotion "family" helps to provide a better understanding of the emotion. Goleman's approach to discrete emotions is also taken in a study where the description of emotions through AD is studied (Igareda & Maiche, 2009). Anger is described as a negative emotion experienced after someone feels responsible for a misdeed or harm to ourselves that results in the will to punish the responsible entity (Smith & Lazarus, 1993). According to Carroll (1990), to feel anger we have to feel that someone has wronged ourselves or people we care for. In the dimensional representations, anger is placed next to fear, as a negative emotion causing high arousal (Lang, 1995).

In the field of psychophysiology, anger is related to an increased heart rate (p. 402) and an increased electrodermal activity (skin conductance response and skin conductance level) (Kreibig, 2010). Such response is related to the results obtained in the study of fear, that share the same pattern (p. 403). As a matter of fact, other scholars present the similarity between these two emotions. Already in the 80s (Ekman et al., 1987), the same similitude was observed and also shared more recently (Stemmler, 2004). As a general overview, on the same level as sadness, fear and anger are the most investigated emotions according to Kreibig (2010). Some scholars have suggested that these two emotions are barely distinguishable (Russell, 1980; Watson & Tellegen, 1985). However, this assumption is made without taking into account emotional behaviour (Adams et al., 2006), since anger is much more related to an approach behaviour in contrast with fear, which is related to avoidance. In this regard, the aforementioned authors suggest that these two emotions are opposed.

4.2.2. Fear

In a discrete approach, fear, as an umbrella term, is related to other terms according to Goleman (2006), such as anxiety, apprehension, nervousness, concern, consternation, misgiving, wariness, qualm, edginess, dread, fright, terror; as a psychopathology, phobia and panic. When located in the framework of emotional dimensions, fear and anger share a lot of characteristics when it comes to physiology. Both fear and anger are considered to be of negative valence and high arousal (Lang, 1995). This dimensional categorisation of fear was already shared in the 80s (Russell, 1980; Watson & Tellegen, 1985). These authors also consider fear to signal threat, assumption that directly relates with the already mentioned contrast between anger and fear in their behavioural response (Adams, Ambady, Macrae, & Kleck, 2006, p. 108). Fear is more related to avoidance than to approach, although it is true that in some cases the feeling of fear can cause the inverse behaviour (see, for instance, Carroll, 1990; or Clasen, 2017, centred on cinema).

One of the most popular comparisons of emotions in the field is anger and fear (Kreibig et al., 2007; Larsen, Berntson, Poehlmann, Ito, & Cacioppo, 2008; Stemmler, 2004). As stated by Kreibig and colleagues, other combinations, such as fear and sadness may be fruitful for contrasting emotions since there were still unexplored combinations. The induction of fear in emotion-elicitation studies has been presented in different forms. Some studies have done so through violent images (Bernat, Patrick, & Benning, 2008), or by showing facial expressions of fear (Lerner, Dahl, Hariri, & Taylor, 2006). Others have done it by means of ordering participants to pose certain facial gestures (Levenson, Ekman, Heider, & Friesen, 1992) or by playing music excerpts conveying fear (Etzel, Johnsen, Dickerson, Tranel, & Adolphs, 2006). In the audiovisual field, short films have been used. In the experiment, such films showed actual violent threat (Montoya, Campos, & Schandry, 2005). For an in-deep review of studies on fear, see Kreibig et al. (2007).

Fear is an emotion that is generally related to broad sympathetic activation. Some trends have been identified by scholars in psychology, namely increased myocardial contractility, vasoconstriction, cardiac acceleration and increased electrodermal activity. Concerning the last two measures, which are used in the experiments found in this project, various studies present increased HR and EDA in single measures when participants are exposed to fear (Kreibig, 2010, p. 404). However, the same scholars point out that trends shown in previous experiences report HR deceleration and also decreased skin conductance level for fear induction (Bradley & Lang, 2000), although the explanation may be found in the fact that the participants were drawn deeper into the feeling of fear, which is characterised by immobilisation rather than an active response (Kreibig, 2010, p. 405). An example of this type of psychophysiological response is the study by Montoya et al. (2005), who induced fear by means of film scenes. They report decreased cardiac output in heart rate for fear, which is contrasted with anger, the second emotion they exposed their participants to. Anger was associated with an increased cardiac output, as already observed in the previous section in this chapter. Therefore, this study shows an actual contrast between the psychophysiological responses derived from these two emotions.

Previous experiences within MA used stimuli portraying fear. Most particularly, Ramos (2015) analysed film clips with AD portraying the emotions of fear and sadness (and also disgust), as they are said to cause changes in the autonomic nervous system for a period of time (Gross & Levenson, 1995).

4.2.3. Sadness

The emotion of sadness can be better conceived when approaching the emotion as a discrete term that covers concepts such as grief, sorrow, cheerlessness, gloom, melancholy, self-pity, loneliness, dejection, despair, and, when pathological, severe depression (Goleman, 2006). When mapping sadness by means of a dimensional approach, in opposition to the previous emotions of anger and fear, it is considered to have negative valence and low arousal (Lang, 1995). Sadness, as stated before, is also one of the emotions that appear more frequently in the study of induction of emotions, as stated by Kreibig et al. (2007). While it has been shown that normally anger and fear are compared in the studies, sadness is compared to happiness. However, there has been a call for more research on further combinations by some scholars (Kreibig et al., 2007).

In their study, Kreibig et al. (2007) list all the different means by which sadness has been induced in experimental conditions. Some previous research experiences include the use of sad pictures (Rainville, Bechara, Naqvi, & Damasio, 2006), the provision of a task in which facial sad expressions have to be posed (Levenson et al., 1992), or the use of sad musical stimuli (Etzel et al., 2006). The induction of sadness through audiovisual material by means of sad films has also been carried out. In the study by Kunzmann & Grühn (2005) newly edited films were aimed at inducing sadness to participants from different ages. Another study (Tsai, Levenson, & Carstensen, 2000) used clips from previously validated films (Gross & Levenson, 1995) to test emotional responses between Chinese and European Americans and across different age ranges. Both studies coincide that physiological responses decrease in adults in comparison to younger participants, as shown in other studies centred on age difference and emotional physiological response (Labouvie-Vief, Lumley, Jain, & Heinze, 2003).

The physiological response to sadness is normally characterised by an overall withdrawal of the sympathetic system. Therefore, sadness is related to a decrease in both heart rate and electrodermal activity. Such results have been found in most part of studies using a variety of stimuli, namely film material, music or imagery (Kreibig, 2010, p. 405). As in the example of fear, where very high levels can lead to immobilisation instead of the expected avoidance behaviour, sadness or “non-crying” sadness has been contrasted with “crying” sadness. As a matter of fact, a handful of studies have reported that when the participant starts crying, measures increase importantly, both for heart rate (Kunzmann, Kupperbusch, & Levenson, 2005; Luminet, Rimé, Bagby, & Taylor, 2004) and electrodermal activity (Kunzmann et al., 2005; Ritz, Steptoe, DeWilde, & Costa, 2000; Tsai, Pole, Levenson, & Muñoz, 2003; Vianna & Tranel, 2006).

The aim in this research is that of eliciting non-crying or deactivating sadness. The materials used as stimuli are not intended to bring the participant to a point where they cry. Therefore, the emotion is considered as low in both valence and arousal axes in the dimensional map and the physiology linked to it will be that of decrease, both in heart rate and electrodermal activity utilised in the experiments presented in previous paragraphs.

Emotion	SAM questionnaire		Psychophysiology	
	Valence	Arousal	EDA	HR
Sadness	Negative	Low	Decrease	Decrease*
Fear	Negative	High	Increase	Increase**
Anger	Negative	High	Increase	Increase
*EDA and HR may increase in crying sadness. **EDA and HR may decrease in extreme fear (immobilisation).				

Table 1. Characteristics of the emotions: dimensional approach and physiological responses

4.2.4. Neutral Emotion

Although it is true that what is referred to as “neutral” emotion is not an emotion *per se*, it is intended to provoke in the participant a response to the absence of emotion. Understood in this way, the “neutral” emotion is used in previous works as the control stimulus, to which the “real” emotions, explored above, are to be compared. In their study, Kreibig et al. (2007) compare the response of participants to the emotions of sadness and fear. Stimuli portraying these emotions are compared to two so-called “neutral” clips, so neutral can be “contrasted” (p. 789) with the two emotions. The clips selected were extracted from a neutral film depicting nature and wild animals, a documentary. Such clips had been previously validated in other studies (Rottenberg et al., 2007). Kreibig and colleagues based their inclusion of neutral films on the experiments carried out by Piferi, Kline, Younger & Lawler (2000). In this experiment, they contrasted the projection of a documentary aquatic film with traditional resting as ways of monitoring the baseline or recovery time, acquiring the former strategy higher level of relaxation.

Following the line of elicitation of emotions through audiovisual clips, Montoya et al. (2005) used psychophysiological measures, particularly cardiodynamic and hemodynamic measures to test the response of the participants to the emotions of fear and anger. A “neutral” emotion was also considered as the “control” emotion and a documentary on the origin of water and its various forms was used to calculate the baseline of the participants for all the measures. This baseline is then taken as a reference to calculate the difference of the values between the neutral film and the emotion films. Thus, it can be argued that the inclusion of a neutral stimulus is recommended when using psychophysiology in studies implementing audiovisual stimuli for the elicitation of concise emotions.

5. Measuring Tools

The focus of this section is the emotional activation when watching films, and the different methodological tools used for its measurement. Lang, considered one of the main scholars in the research of emotions, considers that emotional response can be measured by means three different methods: affective reports, physiological response, and observation of behaviours (Lang, 1969). He suggests that both affective reports and physiological changes offer a relatively easy choice, but technological devices and availability are crucial when conducting experiments. The problem arises when choosing from the vast variety of self-report instruments that exist to measure emotional activation (Bradley & Lang, 1994).

5.1. Self-Report Instruments

Various self-report instruments have been used for measuring emotion in experiments utilising audiovisual stimuli are used (see Table 2). One of the first questionnaires created to measure emotion (particularly mood and mood change) was the Mood Adjective Checklist (MACL) (Nowlis, 1965) based on the rating on a 4 point-scale of 130 adjectives related to different dimensions of mood. Following the same principle, and applied to different pathologies, other questionnaires appeared: examples of these are the Multiple Affect Adjective Checklist (MAACL), and then its revised version MAACL-R (Zuckerman et al., 1983), used to test disorders such as anxiety or depression, translated more recently into Japanese (Yasuda et al., 2003), or the Profile of Mood Stats (POMS) (McNair et al., 1971), applied in the field of psychiatry. By making use of different adjectives to represent four different dimensions for each emotion, the Positive and Negative Affect Schedule (PANAS-X)(Watson & Clark, 1999) expands to different kinds of affects or emotions. This questionnaire was used in Ramos (2015) with blind and visually impaired participants. Within each of the emotions, up to four adjectives were found: for instance, for fear words such as nervous, scared, or terrified, were related. The emotion matched with the scenes selected from previously validated scenes deemed to be emotion-eliciting. This way, the users, by choosing one of the adjectives, were rating the intensity they were experiencing. In her experiment the data proved to be significantly relevant only for the emotions fear and disgust, but not for sadness. As a parallel measure to the self-report instrument, she provided psychophysiological measures with the objective of relating both typologies of measurements.

The questionnaire proposed by Thayer, the Activation-Deactivation Adjective List (AD-ACL) (first published in 1967; see also 1986), is also based on the rating on a 4-point scale of a set of adjectives. It focuses on momentary "activation or arousal states" (Thayer, 1986, p. 1). This questionnaire is considered to be highly related to physiological changes as it focuses on the activation of the arousal and predicts individual-difference variables. Furthermore, it is rapidly administered, approximately in 2 minutes. Another questionnaire is the Affect Grid (Russell et al., 1989), which offers a very simple way to rate the current feelings. It is based on a grid of several squares in which emotions or feelings are written in the corners and mid-points of each margin. Its creators claim that this is consistent with other emotion rating scales, such as the PANAS (presented above). However, as a single-item scale, it augments the possibility of random mistakes and decreases the likelihood of acquiring an overall score.

Another questionnaire that has proved correlation with psychophysiological measures is the UWIST Mood Adjective Checklist, developed by Matthews et al. (1990), which had the aim of refining the inconsistencies or psychometric deficiencies of previous tools for measuring mood and emotion. In their review, Gray & Watson (2007) mention that the UWIST Mood Adjective List questionnaire is being used in the health, performance and work literatures, as well as in other countries. Its complexity, though, might make it unsuitable for the experimental purposes here. Such is the case of the Current Mood Questionnaire (Feldman Barrett & Russell, 1998), whose complexity makes it time-consuming and difficult to use. Bradley and

Lang's proposal of questionnaire (1994), the Self-Assessment Manikin (SAM), uses images instead of adjectives or written language in general by means of schematic drawn avatars that represent the intensity of emotion. As in the case of the Affect Grid, it offers an easy and quick way to measure "affective reaction" (valence, arousal and dominance).

The impact of AD in the emotional state of the user has been tested by Ramos & Rojo (2014). In their study, they based the methodology on what is known as the PACMA theory (Grodal, 2006). PACMA is short for perception, emotion, cognition, and motor action. The theory observes film experience as a flow that goes from perception to real motor action, since "an audiovisual representation can induce different experiences in the audience according to the type of content and form" (Ramos & Rojo, 2014, p. 135). The measurement of emotion activation was performed through a self-report instrument originally conceived for the measurement of the emotional impact of translated versus original texts (Lehr, 2014) that was adapted for audiovisual content. In their results, they corroborated their hypothesis: the creation of AD scripts aimed to provoke an emotional response similar to that of sighted audiences is more difficult to achieve in certain types of films, namely avant-garde or art films. Indeed, results show that emotional responses were more similar for both sighted and visually impaired audiences when measuring emotional response derived from more mainstream productions.

Self-report instruments for the measurement of emotional activation through audiovisual stimuli		
Name	**Author**	**Year**
Adjective Checklist (MACL)	Nowlis	1965
Multiple Affect Adjective Checklist (MAACL)	Zuckerman et al.	1983
Multiple Affect Adjective Checklist - Revision (MAACL - R)	Zuckerman et al.	1983
Japanese MAACL	Yasuda et al.	2003
Profile of Mood Stats (POMS)	McNair et al.	1971
Positive and Negative Affect Schedule (PANAS-X)	Watson and Clark	1999
Activation-Deactivation Adjective List (AD-ACL)	Thayer	1986
Affect Grid	Russell et al.	1989
UWIST Mood Adjective Checklist	Matthews et al.	1990
Current Mood Questionnaire	Feldman Barrett and Russell	1998
Self-Assessment Manikin (SAM)	Bradley and Lang	1994
	Lehr and Schuman	2013

Table 2. Self-report instruments on emotional activation through audiovisual stimuli

5.2. Measuring Physiological Response

There is a vast variety of self-report tools which are used to measure the emotion experienced by the user when exposed to a stimulus. The use of questionnaires, though, is considered to be a subjective manner of approaching the measurement; for instance, they are normally presented to the user before or after the stimulus. The research performed in this field has moved to other approaches, normally combined, which expands to the broader area of user experience (UX). Changes in the physiology of the organism have proved to relate to what we feel and experience. The recording of such biological markers is considered a more objective way of quantifying concepts such as emotion.

One method used in the study of the reactions of the organism is the employment of psychophysiological measures. This kind of technology, applicable to a wide range of scientific specialities, allows the recording of indicators of arousal such as brainwaves, muscle contraction, electrodermal activity or heart rate. The application of psychophysiology to the measurement of the organism's arousal can be found in very different stimuli which differ both in the measures used, the stimuli presented and the participants taking part in the experiments. This section is devoted to this multiplicity in the use of such technologies.

The possibilities of this type of measurement allowed for a variety of applications in different contexts. For instance, the measuring method fMRI (functional magnetic resonance imagining), based on the oxygenation patterns across the areas of the brain, is used in the study by Lee et al. (2004) on spatial presence in virtual environments. The measurement of brain activity is also performed in Mikropoulos et al. (2004) while the participant is exposed to a virtual environment. In this case, the electrical current of the brain is measured through electroencephalography (EEG). Postural position is used as measurement of presence. The alteration of the body position is recorded in different studies (see, for instance, Freeman et al., 2000; Regenbrecht et al., 1998; Usoh et al., 1999).

There are some examples in which psychophysiology has been applied to the study of emotions. Such is the case of the study by Baumgartner et al. (2006), where the stimuli used were video and music excerpts targeted at the elicitation of emotions. In the same vein, skin conductance was used to measure the elicitation through the use of music, for the purpose of clarification of musical emotions (Khalfa et al., 2002). Considered as less immersive than audiovisual products, static images are used in a study involving EEG (electroencephalogram) and the effects provoked by the colour red (Kuniecki et al., 2015) or, accompanied by certain auditory stimulus, static images were also used during the measurement of the evoked cardiac response (ECR)[6] (Kuniecki et al., 2003). Images were also used in Bradley & Lang (1994) to test emotions

[6]The evoked cardiac response (ECR) may be described as the sum of two independent response components: an initial HR deceleration (ECR1) and a slightly later acceleration (ECR2) hypothesized to reflect stimulus registration and cognitive processing load, respectively. (Lawrence & Barry, 2009)

elicitation and memory. In this case, the measure was pupil diameter, which was recorded by means of eye-tracking technology.

Different subjects were involved in the experiments by Daltrozzo et al. (2010). They measured skin conductance responses in low-responsive patients, almost all of them in coma, through emotional and neutral stimuli for the recording of their reactivity. Although all these investigations came from fields such as psychophysiology itself or neuroscience, some studies apply those methods to other fields such as interpreting, situated closer to the framework in which this study takes place. Kurz (2003) compares the stress undergone by novice and expert interpreters provoked by the complex circumstances in which they work utilising “pulse rate” and “skin conductance level”, as she names them, when doing simultaneous interpreting.

Such methodologies are found in experiments with audiovisual stimuli. A good example is in the field of videogames, a field that has attracted the use of psychophysiology due to a main feature which make videogames different from “typical” audiovisual products: interactivity. O’Hagan (2016) takes into consideration physiological data based on previous experiences (see Albert & Tullis, 2013) in order to measure the UX, which is defined by ISO 9241-210 (ISO, 2010) as “a person’s perceptions and responses that result from the use or anticipated use of a product, system or service”. O’Hagan measured both HR and EDA (2016) to test the gamer emotions derived from the use of a videogame. Her results were not conclusive even if some of the peaks correlated with the stimuli provided. The author herself states that the stimuli used (based on humour) might not have been strong enough to trigger such emotion markers. It is worth mentioning that one might signal interactivity as one of the key differences between a videogame and a traditional cinematic product, augmenting the immersion in the former. However, there have been some experiences with the use of more “traditional” audiovisual stimuli, even if they are considered less immersive.

Wilson & Sasse (2006) used HR, EDA and blood volume pulse (BVP), in opposition to more subjective and traditional methods when testing the quality of image according to frame segmentation in video stimuli. Physiological responses were used to test interactivity and entertainment technologies in Mandryk, Inkpen & Calvert (2006), showing, for example, how this can change when the user plays against the machine or against another person. Moving to the measurement of the response generated by cinema itself, Baños et al. (2004) relate the affective content of audiovisual materials with the experimental factors of immersion and presence. Dillon (2006) also employs cinematic stimuli in her study. She collected data from psychophysiological markers – electrodermal activity and heart rate – to test presence when exposing the participant to emotional and non-emotional arousing films. More recently, electroencephalography (EEG) was proposed in the study by Kruger et al. (2016) about immersion in subtitled content. They aim to complement subjective measures (questionnaires) with the collection of physiological data (beta coherence between the prefrontal and posterior regions of the brain), deemed to be more objective.

The physiological markers that can be studied are vast. Physiology has proved to be a group of phenomena that is complex to measure. As stated by Kreibig (2010, p. 410), "non-emotional physical, behavioural, and psychological factors affect psychological activation before, during, and after emotion". That is why the use of more than one response is studied at the same time. The same author argues that for the study of the response of the ANS to emotions the two indicators that are most used in combination are HR and EDA (23 in her analysis of 137 studies) (p. 409).

6. From Questionnaires to Physiology in AD and AST Research

The analysis of user experience when exposed to AST and AD has previously been tested, although to a very low extent in AST. Examples of this are the experiments carried out by Thrane (2013), who tested the delivery of synthetic AST in a pilot study for the Swedish channel DR1Syn with 16 blind and partially sighted participants, and by Holsanova et al. (2015), who tested the delivery of AD and AST soundtracks for TV through smartphones on 11 participants (some of them blind or partially sighted) also in Sweden. They both based their results on the use of questionnaires on preferences.

The use of self-report questionnaires is the most common approach when studying these two services. Since studies on AST are not very numerous in research, it is relevant to mention some of the works carried out on users' reception and preference of AD. In his PhD thesis, Cabeza-Cáceres (2013) focuses on intonation, speed rate and explicitness of the AD and their effect on comprehension. Szarkowska & Jankowska (2012), for example, discuss the application of synthetic voices for the narration of the AD and run some experiments in which human voices were compared to TTS voices. This concept is close to that of the more recent study by Fernández-Torné & Matamala (2016), where synthetic and human voices in AD were compared in films dubbed into Catalan. The reception of TTS voices was also studied in AST (Szarkowska & Mączyńska, 2011) and also applied to a specific genre of audiovisual products, audio described films for children (Szarkowska & Walczak, 2012; Walczak, 2010).

However, studies on reception of AD by means of questionnaires focusing have seen some limitations, and new approaches are recommended. Chmiel & Mazur (2016) consider this type of questionnaires as too subjective and introspective. The results of their study on user's preferences (about different AD solutions) proved to be largely inconclusive and they recommended for future work on the cognitive efficiency of AD, that is, obtaining the maximum benefit with the minimum effort, allowing for the creation of mental models.

Whereas no studies on the measure of flow are found on AD or AST, presence in AD has been studied by some scholars. Freeman and Fryer (Fryer, 2013; Fryer & Freeman, 2013b) state that a good quality AD is that which passes unnoticed, this leading to the assumption that the user feels transported and immersed into a mediated reality, thus, developing a feeling of non-mediation (Lombard & Ditton, 1997). Fryer and Freeman (2013), both having experience in the way AD is perceived (see Freeman, 2010; Fryer, 2013; Lessiter, Freeman, Keogh, & Davidoff, 2001; Walczak & Fryer, 2018) provide a comparison between the levels of presence in blind, partially sighted and sighted audiences when exposed to audiovisual media by using specific questionnaires (see Section 5.1.).

Their experiment on presence allowed for the rating of the quality of AD "not by focusing on the description itself, but on the user's experience of the source programme, film or play" (p. 1), that is, by measuring the levels of presence. In the article, they argue that not only the presence experienced by people with different sight conditions can be compared but also that it can be done with sighted audiences. There is proof that, in some cases, blind and partially sighted audiences are able to experience transportation more than their sighted counterparts when exposed to certain types of AD. In a recent study, Wilken & Kruger (2016) take into account immersion and the feeling of being transported when the audience is exposed to different kinds of AD. They considered *mise-en-shot* elements when modifying the AD scripts that were tested with blind participants by means of self-report questionnaires. The results showed that the exclusion of *mise-en-shot* items caused a dip in immersion levels. The results for the quantitative data in terms of transportation proved to be statistically irrelevant. However, the qualitative data showed some differences. It has to be said that the experiment by Wilken & Kruger was performed with sighted audiences as a foundational experiment before running tests with blind participants.

Emotional elicitation is in turn measured by means of questionnaires in Ramos (2015), as it has been discussed previously in this chapter, or in Ramos & Rojo (2014), where the self-report questionnaire developed by Lehr and Schuman (Lehr, 2014), a priori created for the emotional impact of original and translated texts, was altered to fit the requirements of their study.

Although questionnaires have been used to a great extent, not many scholars have introduced psychophysiological measurements to the study of reception of audiovisual media by blind and partially sighted audiences. AD has been tested in sighted audiences by means of eye-tracking (Igareda & Maiche, 2009) in order to make an impact on the quality of AD identifying the elements causing emotion in a film. Visual elements and sounds were tested also with eye-tracking for the purposes of AD elaboration (di Giovanni, 2014; Krejtz et al., 2012; Orero & Vilaró, 2012; Vilaró et al., 2012). Yet, experiments run with actual blind and partially sighted audiences, primary receivers and users of AD, using psychophysiological markers have only been performed, to the best of our knowledge, by two authors. In their studies, HR (Ramos, 2015) and HR and EDA (or galvanic skin response) (Fryer, 2013b) have been used alongside questionnaires.

Ramos (2015) includes psychophysiological measures to the study of AD reception. She exposed the participants to clips corresponding to the emotions of fear, disgust and sadness with a duration between 3 and 4 minutes. These clips had been validated in previous studies (Fernández Megías et al., 2011). The scenes were elicitors of emotions fear, disgust and sadness. According to the answers in the questionnaires, she concluded that the emotion elicitation was particularly remarkable for the emotion of fear and that the scenes portraying sadness, which were accompanied with a certain type of music and dialogue, did not necessarily rely on the AD. As far as the psychological measures are concerned, she measured the baseline of the heart rate before the clip was started and then compared it to the average acquired during the clip. The results, though, fell short from what was expected. The data collected did not correlate with the results in the questionnaires, even though such emotions are deemed to conquer the autonomic nervous system. However, the measurement of heart rate does provide significant results when comparing the different emotions (fear, disgust and sadness). This leads to the assumption that the analysis of such physiological response is worth the effort and can generate interesting results. She signals that the autonomic nervous system is also subject to other processes of the body, such as muscle demands or digestion. Closing the discussion of physiological results, she calls for more research on the topic and advises to complement such a measure (HR) with other measuring methods (EDA, heart rate variability [HRV] and cortisol levels).

Fryer (2013), on the other hand, analyses physiological responses as an extra part of her main experiment. She concludes that such measurements do not provide enough scientific relevance. Physiological data collected did not follow any consistent patterns and HRV was the only measure that proved significant. It is true, though, that fear was the most eliciting emotion in general. The author mentions two different explanations that somehow could have influenced somehow negatively the outcome of the experiments. First, the researcher herself is a professional audio describer, and participants recognised her voice in the human-voiced AD used in the stimuli. Secondly, the device used was a custom-made polygraph, with a custom-made software that may not have been accurate enough.

7. Conclusion

According to the theories presented in this chapter, audiovisual content is created and consumed with the principal objective of entertainment. The relationship between audience and media has been the object of research in mass communication and media psychology (Jennings Bryant & Miron, 2002; Miron, 2006; Vorderer & Hartmann, 2008; Zillman, 1999; Zillmann & Bryant, 1994). The main gratification extracted from filmic material is the pleasurable effect distilled from entertainment (Vorderer & Hartmann, 2008). Pleasure is acquired by practising the continuous balancing of emotions that allows the audience to even experience entertainment when faced with what are considered to be negative emotions, such as pain, sadness or fear (Carroll, 1990; Rozin, Guillot, Fincher, Rozin, & Tsukayama, 2013; Schramm & Wirth, 2010).

Such entertainment is fostered when feeling involved in the content (Lyons et al., 2014). Here the explanation has been sought in the theories of flow and presence. The theory of flow was first described by (Csikszentmihalyi & Czikszentmihalyi, 1988) as a feeling that can be developed in a user when exposed to any kind of task. This task should present a series of characteristics that would allow for a balance between the user's skills and the difficulty posed by the task, in this case, filmic materials. Presence describes the feeling of being there, of experiencing elements other than the ones present in the real space. Presence has been studied in cinema, TV or videogames, as well as in books.

Emotions have also been shown to be the essence of entertainment and the pleasurable experience sought in audiovisual consumption. Entertainment, as the big umbrella embracing the filmic experience, then, finds its meaning in the generation and balancing of emotions (Vorderer & Hartmann, 2008, pp. 540-545); and such link is found with presence (Freeman et al., 2000) and flow (see, for instance, Grodal, 2006, p. 4).

Emotions have been the object of fruitful research: there are many theories for their definition and categorisation and it is important for the purpose of each study to delimit and define what is the approach selected. Essentially, emotions can be categorised by means of two different approaches: discrete and dimensional. The theoretical framework here is based on the hybridisation of this duality that has proved to better understand and shape both emotions and emotional responses both at a theoretical and practical level. It conceives them as units with their own intrinsic characteristics which also serve as umbrella terms embracing other nuances. These hypernyms, as well as its different nuances are, in turn, placed along dimensions that move in a spectrum. Most accurately, and despite the existence of different emotions, the two axes present in our map will be that of valence and arousal, common to all studies using the dimensional approach and applying psychophysiological measures.

For the measurement of emotions different methodologies have been used, not only in the field of psychology but also in a variety of scientific fields. First of all, and by following a more subjective approach, many scholars have worked on self-report instruments. Most part of these self-report instruments have been created to be applied to experiments of virtual environments, however others are said to be able to embrace a more heterogeneous group of stimuli. Such questionnaires, though, are considered subjective and some scholars propose or have already used mixed approaches in which questionnaires are accompanied by psychophysiological measures (see Chmiel & Mazur, 2016; Fryer, 2013; O’Hagan, 2016; Ramos, 2015; Wilken & Kruger, 2016).

The way our organism changes in front of a stimulus or a group of stimuli would, theoretically, constitute a more reliable and objective way of recording and quantifying the experience lived by the user. Such methodology was used in two studies with blind and partially sighted participants and AD stimuli (Fryer, 2013b; Ramos, 2015). Physiological data in the two studies did not provide very clear results that were comparable to self-report responses. However, both authors call for more investigation with the use of such measures. More accurately, the use of other measures is recommended to complement the more “traditional” self-report instruments. The application of such methodology counts with very little research in the field of audiovisuals and user experience in general. Its outcomes have proved relevant and noticeable in other studies in psychology and more specifically, in the field of emotion research.

Chapter 4

Study 1: Descriptive Analysis

Chapter 4

Methodology of Study 1: Descriptive Analysis

This chapter aims to present the methodology of Study 1. This study is composed of the descriptive analysis of a corpus of 6 multilingual audio described films. The focus is put on audio subtitles that are integrated, human-voiced and broadcast-mixed, following the terminology proposed in Chapter 2. The analysis takes into consideration the AST in six films, all multilingual and produced in the 21st century. As already explained in Chapter 2, some scholars have already carried descriptive research on AST focusing on dubbed films in German (Benecke, 2012), multilingual audio described and audio subtitled films (Braun & Orero, 2010; Harrouet, 2016), and Dutch multilingual films (Remael, 2014).

Section 1 details the selection of materials for the corpus of films and **Section 2** moves to the method of analysis. Results are presented and discussed in Chapter 5.

1. Materials: Corpus Selection

Since the aim was to analyse audio subtitled films, it was considered that multilingual films would be suitable materials. The time-frame was also limited to 21st century films in order to analyse more updated practices, and the format chosen was DVD. English was selected as the main language in the multilingual films to be considered because the choice of audio subtitled content was expected to be wider. A last selection criteria was that all films should include the specific tag “audio description” on the online platform Amazon.

Following the previous criteria, a search process began, using available materials in the workplace, personal knowledge and articles from where this analysis departs (Benecke, 2012). Other sources were consulted such as the database created in the TraFilm project (TraFilm, 2017), which allows one to navigate across a collection of multilingual films and scenes; studies such as Bartoll (2006) and Bréan & Cornu (2012), and a variety of forums specialised in cinema or languages (IMBd, Polyglot Nerd, Wikipedia).

A total of 20 films were identified. The first step was a preliminary visualisation of the original version of all these films, to understand the number of contributions in secondary languages in the dialogues and therefore the possibility to include audio subtitles. The original version was prioritised as it was easy to obtain through local libraries without the need to actually purchase the DVD addressed to the English market that included the AD and the AST. The first selection of films included:

A Good Year (Scott, 2006)
Memoirs of a Geisha (Marshall, 2005)
The Science of Sleep (Gondry, 2006)
Syriana (Caghan, 2005)
Borat: Cultural Learnings of America for Make Benefit Glorious Nation of Kazakhstan (Charles, 2006) (from now on: Borat)
Saawariya (Bhansali, 2007)
Fury (Ayer, 2014)
To Rome with love (Allen, 2012)
Argo (Affleck, 2012)
Midnight in Paris (Allen, 2011)
Salt (Noyce, 2010)
My Name Is Khan (Johar, 2010)
Grand Torino (Eastwood, 2008)
Valkyrie (Singer, 2008)
Slumdog Millionaire (Boyle & Tandan, 2008)
Eat Pray Love (Murphy, 2010)
Brick Lane (Gavron, 2007)
Bride & Prejudice (Chadha, 2004)
Slumdog Millionaire (Boyle & Tandan, 2008)

After the preliminary visualisation, some contents were discarded since not all them were appropriate for an analysis of AST. Some films do include scenes in secondary languages but these are used as a “decorative” feature of the dialogue, being irrelevant to the understanding of the contents. As indicated before, the visualisation of the original versions of the films without the AD/AST service was performed with the sole purpose of deciding if the film should be acquired in the original version which included the AD track in English. The films discarded were:

A Good Year (Scott, 2006)
Memoirs of a Geisha (Marshall, 2005)
Eat Pray Love (Marshall, 2005)
Valkyrie (Singer, 2008)
Gran Torino (Eastwood, 2008)
Fury (Ayer, 2014)
To Rome with Love (Allen, 2012)
Argo (Affleck, 2012)
Midnight in Paris (Allen, 2011)
Salt (Noyce, 2010)
Saawariya (Bhansali, 2007)
Brick Lane (Gavron, 2007)
My Name Is Khan (Johar, 2010)

A Good Year is all in English and only presents some interventions in French; the film *Memoirs of a Geisha* presents solely a few expressions in Japanese; and in *Eat Pray Love*, Julia Roberts offers some words and sentences in Italian. Likewise, the film *Valkyrie*, identified as a multilingual film, as English and German featured, has to be discarded as only one initial contribution is uttered in German. Other films such as *Midnight in Paris* and *From Rome with Love*, by Allen, present dialogues entirely in English even if they are set in France and Italy respectively. In *Gran Torino*, the dialogues in the Hmong language are very reduced and have not been translated in the form of subtitles. The same happens with *Argo* and its dialogues in Arabic and Persian. *Salt*, with Angelina Jolie as main character, presents few sentences scattered during the film in Russian and Korean. The Bengali in *Brick Lane* is very little, since all its main characters who are supposed to be Bangladeshi, speak in English in the original version.

On the other hand, the films *Saawariya* and *My Name Is Khan* were discarded. In both cases, the language unknown to the English-speaking audience, Hindi in the case of *Saawariya,* and Hindi and Urdu in the case of *My Name is Khan,* are the main languages of the films and English is used only occasionally. Therefore, the films are mainly monolingual. The final list of films analysed is the following:

1. *Everything Is Illuminated* (Schreib, 2005)
2. *The Science of Sleep* (Gondry, 2006)
3. *Syriana* (Caghan, 2005)
4. *Borat* (Charles, 2006)
5. *Slumdog Millionaire* (Boyle & Tandan, 2008)
6. *Munich* (Spielberg, 2005)

The next step was to acquire the DVD with AD and AST addressed to an English-speaking market. Some general information concerning the languages involved in the film and the presence of AST are commented as follows:

Everything is Illuminated (Table 3) features English as the main language and dialogues in Russian and Romanian. This film contains dialogues in two secondary languages.

Everything is illuminated (2005)	
Run time	106 min
Nationality	USA
Languages	English, Russian, Ukrainian
Characters	Father Grandfather Alex Man roadside Brother Boy with goats Waitress Listah Well digger 1 Well digger 2 Well digger 3

Table 3. General information about *Everything Is Illuminated*

The Science of Sleep (Table 4) contains three different languages: English, French and Spanish. Spanish is only present in two expressions uttered by the protagonist, Gael García, who is Mexican. French, though, is spoken a lot and the film, which is set in France. Trilingual Stéphane shows his preference towards talking in English because he feels more confident.

The Science of Sleep (2006)	
Run time	106 min
Nationality	USA
Languages	French, English, Spanish
Characters	Stéphane Stéphanie Zoé Christine Guy Serge Martine Mr Pouchet Gérard

Table 4. General information about *The Science of Sleep*

Syriana (Table 5), an international thriller, which presents the economic and war interest in the current interaction between western and eastern countries. Although English is, undoubtedly, the main language of the film, other languages such as Farsi, Urdu or Arabic are spoken in different scenes. In some cases the languages spoken are also the topic of the scene. Such is the case in a scene where an Arabic-speaking man tells a young Farsi-speaker to learn Arabic in order to get a job in the port as fisher.

Syriana (2005)	
Run time	122 min
Nationality	USA
Languages	English, Urdu, Arabic, Farsi, French, Mandarin
Characters	Bob Arash Wasim Kahn Salim (Wasim's father) Emir Farooq Prince Nasir Al-Subaai Prince Meshal Al-Subaai Man at the port Soldier Koran teacher Sayid Mohammed Sheik Agiza Man after torture Mussawii Supplicant Young man on video Nasir's man (debate) Cricket player Radio speaker (male voice)

Table 5. General information about *Syriana*

The film *Borat* (Table 6) presents English as the main language of the film and interventions in a lot of languages. As a matter of fact, most of the time the secondary language spoken is just an invented language out of the combination of many Eastern European languages and Hebrew.

Borat (2006)	
Run time	80 min
Nationality	USA
Languages	English, Romanian, Hebrew, Polish, Armenian
Characters	Borat Azamat Borat's wife Neighbour

Table 6. General information about *Borat*

Munich (Table 7) is a film based on the Israeli operation "Wrath of God" against Palestinian Liberation Movement. Its dialogues are mainly in English combined with Arabic and Hebrew and other languages such as German, French, Italian, Greek or Russian.

Munich (2006)	
Run time	164 min
Nationality	USA
Languages	English, Arabic, Hebrew, German, French, Italian, Greek, Russian
Characters	Avner Andreas Yvonne Wael Papa

Table 7. General information about *Munich*

Slumdog Millionaire (Table 8), based on the story of a boy raised in the suburbs of Mumbai, features English and Hindi. The presence of multilingualism in this film is basically divided temporally. Almost all dialogues are in Hindi when the main characters are kids and when showing present time, where the characters are adult, the dialogues are mostly in English among them.

Slumdog millionaire (2008)	
Run time	123 min
Nationality	United Kingdom
Languages	English, Hindi
Characters	Jamal Salim Mother Latikah Prakash Man on rooftop Man shouting (fight) Maman Punnoose Bearded Man Policemen

Table 8. General information about *Slumdog Millionaire*

2. Methodology: Data Extraction and Analysis

Based on a literature review (Benecke, 2012; Braun & Orero, 2010; Harrouet, 2016; Matamala, 2014; Remael, 2012) and a first visualisation of all films, the most relevant features to be analysed were selected. These included: the number and type of voices, the audibility of original and AST tracks, the type of isochrony of the audio subtitles, the prosody, the assignation of characters speaking, the AST effect, and the textual features of the audio subtitles in comparison to written subtitles. For each of the features, different possibilities were identified. These possibilities were defined and coded as explained next.

The first item is the identification of the **voices** used to deliver the AST and AD track of the film. The following options were considered:

- The same male voice for the AD and the AST (Only male AD/AST)
- The same female voice for the AD and the AST (Only female AD/AST)
- A male voice for the AD and a female voice for the AST (Male AD/Female AST.

- A female voice for the AD and a male voice for the AST (Female AD/Male AST)
- A female voice for the AD and female characters, and a male voice for male characters (Female AD + AST female characters/Male AST characters)
- A male voice for the AD and male characters, and a female voice for female characters (Male AD + AST male characters/Female AST characters)

Differences in the **audibility** of the two audio inputs in an audio subtitled scene can play a crucial factor in the representation of multilingualism for people who cannot access the visuals. The contrast between the volumes is made by making the distinction between:

- The original can be heard (Heard)
- The original cannot be heard (Not heard)

It has to be said, though, that the differences in the audibility of both tracks has not been tested in terms of decibels at a quantitative level due to the technical complexity when splitting the tracks, since they have been merged with the original in a single track. The analysis is therefore based on the subjective perception of the researcher.

Isochrony in this analysis is referred to the synchronisation between the duration of the AST and the original line. Thus, the delay on the AST's behalf in relation to its corresponding line in the original dialogue will have different forms when audio subtitling a scene. Whereas some AST may synchronise perfectly with the corresponding original line, others are delivered before or after the original line or are overlapped for some period of time. According to Sepielak (2016, p. 1060), who approaches synchronisation in voice-over, "it is undefined whether the gap between the start of the original and the onset of translation should be calculated in seconds, milliseconds, frames, or audible words". She differentiates three main sorts of isochrony: full, where at least one word of the original is heard before and after the translation; initial, when one word is heard before; and final, when a word of the original is heard after the translation. For this descriptive analysis Sepielak's terminology is used and adapted to the needs of AST. The classification goes as follows:

- Before, when the whole AST is delivered before the original line;
- After, when it is delivered once the original is finished;
- Later, when the AST comes in little delay and is overlapped;
- Synchrony, when there is full synchronisation between original utterance and AST, that is, they start and finish at the same time;
- Initial, when both original and AST begin at the same time but the ending is different;
- Final, when the start is different but there is synchronisation at the end of both the AST and the original utterance.

Assignation is understood as the verbal explanation on the audio describer's behalf of who is going to speak and the way the speech will be uttered. As a general idea, it is closely related to the notions of reported and direct speech. That is, for instance, the existence of "he/she says" expressions concerning an AST. The role of the AD is very helpful in cases where only one voice is used to deliver the AST for more than one character in a scene. The distinction in assignation is made according to the criterion:

- Direct, the assignation of the characters relies on the delivery of the AST.
- Reported, there are indications from the audio describer.

The final aspect, and one of the most relevant ones concerning character recognition is **prosody**. Matter-of-factly, acting AST is of great help when distinguishing who is intervening in a scene, since, as mentioned before, not all productions count on more than one or two (male and female) voices to carry out the process to make a film accessible for the blind and the partially sighted. Nevertheless, some productions rely on the audibility of the original lines to provide the user with the actor or actress' performance, whereas others tend to recreate the performance in the AST. The distinction is made according to the theoretical explanation provided in Chapter 2, on the prosodic features of the dubbing and the voice-over effect. Audio subtitles are categorised in terms of:

- Read, with reading aloud features.
- Acted, imitating prosodic features of the fake orality found in dubbing.

The complexity of this service, which takes an oral text in a foreign language, then translates it into a written subtitle in the target language and then transfers the written text to an oral text again in the target language, can also produce changes in the characteristics of the text that carry information about the oral or written nature of the text (to be read or to be heard). The **textual features** of the AST service are commented under this label. While this information might not be relevant to people unable to read subtitles, such as the blind and the partially sighted, it is of crucial importance when considering people with reading impairments, who may be able to partially read and hear at the same time the written subtitle and the corresponding audio subtitle. The distinction here is made between

- Verbatim, when there are no changes
- Changed, when the AST are different from the written subtitles through omission or expansion.

A first data extraction procedure was tested on the film *Everything is illuminated* using a Word table, but it proved challenging to analyse the data accurately. This is why the different items and assigned values were then translated into an Excel file (see "AST Scene Analysis" in the Electronic Annexes), with the following fields:

1. Scene number
2. Initial time code of the scene
3. Final time code of the scene
4. Voices

5. Prosody
6. Assignation
7. Character who utters the subtitle/audio subtitle
8. Transcription of the subtitle/audio subtitle
9. Isochrony
10. Audibility of the original
11. Effect
12. Textual features
13. Observations that might exist

This Excel was used to carry out a preliminary analysis of the first film, *Everything Is Illuminated*. While most part of features were gathered at scene level, isochrony was gathered at utterance level, since changes were observed within the same scene along audio subtitles. The proposed analysis and tools were considered fit for purpose and no changes were made to the Excel file.

Then, a second visualisation was performed to fill in the data for each film in the Excel file. The software used for visualising the films was VLC Player. It is important to state the software since the time marks vary from software to software, and it can affect the spotting of scenes.

The data extracted in the Excel file were the basis for the analysis presented in the next chapter. For each film, a so-called matrix following the structure shown on Table 9 was created in which the main trends are summarised. For each feature, the main trend is included. While most of the features had the same characteristics along the whole film (and this has been included in the matrix), isochrony varied along audio subtitles. In order to establish a trend of the type of isochrony predominant along each of the films, a transcription is provided of each audio subtitle, which is then tagged with the labels concerning the different types of isochrony (see “AST Scene Analysis” in the Electronic Annexes). Afterwards, percentages are extracted in relation to the total number of audio subtitles in each of the films, and the trend is reported.

Title	
Audio subtitled scenes	
Audio subtitled time	
Audio subtitles	
Voices	
Assignation	
Audibility	
Prosody	
Isochrony	
Effect	
Textual features	
Observations	

Table 9. Summary matrix for the results of Study 1

Also, the data included in the Excel were used to provide a more thorough analysis of each category, with a comparison across films. As already indicated, the approach was mainly qualitative and aimed to provide an overall vision of the main trends in the production of AST without entering into in-depth quantitative details. This is especially relevant taken into account the limited bibliography on the topic of AST. It has also been useful when creating the stimuli for the experimental part of the thesis.

Chapter 5

Discussion of Results in Study 1

Chapter 5
Discussion of Results in Study 1

This chapter discusses the results of the study described in Chapter 4. In **Section 1**, the matrixes of analysis for each of the film are presented, including the categories identified in the previous chapter: title of the film, number of scenes audio subtitled and total time of such scenes, total number of audio subtitles, voices, assignation of characters, audibility of the original and the AST track, and the prosodic features in the delivery of AST. Out of these features, the matrix states if the overall effect used in the film is dubbing or voice-over. Finally, information is also provided about the textual features, that is, if they are read verbatim or some expansion/omission has taken place. **Section 2** moves to a discussion of the main results. In this section more in detail information for each of the features analysed is provided. The information is organised according to the structure of the matrix of analysis.

1. Matrix of Analysis: General AST Characteristics

First, the tables for all the films are presented so as the reader can schematically scan through the outcome for each film: *The Science of Sleep* (Table 10), *Everything Is Illuminated* (Table 11), *Syriana* (Table 12), *Slumdog Millionaire* (Table 13), *Borat* (Table 14) and *Munich* (Table 15). A final table (Table 16) summarises all the films and all the features. Afterwards, there is a discussion section where the results are commented on following the order of the features presented in the matrix of analysis.

The Science of Sleep	
Audio subtitled scenes	36
Audio subtitled time	26 min 33 sec
Audio subtitles	384
Voices	Female AD + AST female characters/Male AST characters
Assignation	Direct
Audibility	Not heard
Prosody	Acted
Isochrony	Before (42%)
Effect	Dubbing
Textual features	Verbatim
Observations	
The film is surrealist and abstract, thus incrementing visual multiplicity (Maszerowska, 2014). That means sometimes we move in and out of the characters' mind, voices are the same but sometimes AD cannot express the division of space.	

Table 10. Details of *The Science of Sleep*

Everything Is Illuminated	
Audio subtitled scenes	26
Audio subtitled time	25 min 58 sec
Audio subtitles	318
Voices	Female AD + AST female characters/Male characters
Assignation	Direct
Audibility	Original heard
Prosody	Acted
Isochrony	Later (37%)
Effect	Voice-over
Textual features	Verbatim

Table 11. Details of *Everything Is Illuminated*

Syriana	
Audio subtitled scenes	25
Audio subtitled time	19 min 1 sec
Audio subtitles	250
Voices	Female AD + AST female characters/ Male AST characters
Assignation	Direct
Audibility	Original heard
Prosody	Acted
Isochrony	Later (43%)
Effect	Voice-over
Textual features	Verbatim

Table 12. Details of *Syriana*

Slumdog Millionaire	
Audio subtitled scenes	27
Audio subtitled time	17 min 45 sec
Audio subtitles	200
Voices	Male AD + AST male characters/ Female AST characters
Assignation	Reported
Audibility	Original heard
Prosody	Read
Isochrony	After (49%)
Effect	Voice-over
AST = written subtitle	Verbatim*
Observations	
(1) Only one written subtitle “We won’t” and AST “We don’t” (Scene 1. 5:53-6:30) *Written subtitles in the original version have different colours when displayed on screen, this effect is not simulated by using different male voices in the AD version.	

Table 13. Details of *Slumdog Millionaire*

Borat	
Audio subtitled scenes	32
Audio subtitled time	8 min 7 sec
Audio subtitles	138
Voices	Male AD + AST male characters/ Female AST characters
Assignation	Reported (1)
Audibility	Original is heard
Prosody	Acted
Isochrony	Later (42%)
Effect	Voice-over
Textual features	Verbatim
Observation	
(1) AD reports speech in some scenes (Scene 29.53:11-54:50) or Scene 31 (1:08:55-1:09:28)	

Table 14. Details of *Borat*

Munich	
Audio subtitled scenes	5
Audio subtitled time	3 min 23 sec
Audio subtitles	45
Voices	Only female AD/AST
Assignation	Reported
Effect	Voice-over
Audibility	Original is heard
Prosody	Read
Isochrony	Before (40%)
Textual features	Verbatim

Table 15. Details of *Munich*

Before discussing each of the features analysed in the corpus, in Table 16 the summary of the main trends found in the corpus is found. In the first row, all the different features analysed are shown. In the first column, the title of each of the films in the corpus is stated.

Features / Film	Voices	Assignation	Effect	Audibility	Prosody	Isochrony
The Science of Sleep	Female AD + AST female characters/Male AST characters	Direct	Dubbing	Not heard	Acted	Before (42%)
Everything Is Illuminated	Female AD + AST female characters/Male characters	Direct	VO	Heard	Acted	Later (37%)
Syriana	Female AD + AST female characters/ Male AST characters	Direct	VO	Heard	Acted	Later (43%)
Slumdog Millionaire	Male AD + AST male characters/ Female AST characters	Reported	VO	Heard	Read	After (49%)
Borat	Male AD + AST male characters/ Female AST characters	Reported	VO	Heard	Acted	Later (42%)
Munich	Only female AD/AST	Reported	VO	Heard	Read	Before (40%)

Table 16. Summary of the features in the matrix of analysis.

2. Discussion

This discussion comments first on the presence of AST in the films analysed. It moves later to discuss the aspects relevant to the form and delivery of audio subtitles: voices, assignation, audibility, isochrony and prosody. From the conclusions extracted from each one of these items, the discussion moves to the distinction between dubbing and voice-over effect in order to identify the main trend in the films analysed.

Secondly, the textual characteristics are also discussed: the comparison between written subtitle and AST and the translation features used (if they have been expanded or if there are omissions).

2.1. Presence of Audio Subtitles

In general, and as expected, audio subtitled scenes are a minority in the films analysed. As multilingual films, only secondary languages are offered with subtitles that are then adapted into audio subtitles. Out of the six films in our corpus, the film with more audio subtitled time is *The Science of Sleep* with 36 scenes audio subtitled and a total of 384 audio subtitles in 26 min and 33 seconds (25.04% of the film contains AST) featuring French and only 2 originals in Spanish.

This film is followed by *Everything is Illuminated,* which has 26 audio subtitled scenes and a total of 318 audio subtitles with a duration of 25 min and 58 sec (24.49% with AST). *Syriana,* which presents 25 scenes with a total of 250 audio subtitles, from Arabic, Urdu, Farsi and French, appearing along 19 min and 1 second (15.58% with AST).

The film *Slumdog Millionaire* presents 27 scenes with audio subtitles, with a number of 200 audio subtitles and 17 min and 45 sec of audio subtitled content (14.43% audio subtitled content). In *Borat* and *Munich*, the presence of scenes containing secondary languages is much more occasional. *Borat* has 32 short audio subtitled scenes with 138 audio subtitles along 8 min and 7 sec (10.13% of audio subtitled content). In *Munich* only 5 scenes are audio subtitled with a total of 45 audio subtitles in 3 min and 23 sec (3.25% of audio subtitled content).

2.2. Voices

All the films analysed present two different voice talents in the film. There is always a male and a female voice. The male or female voice delivering the audio description is also in charge of providing the audio subtitles for male or female characters, respectively. For the six films analysed in the corpus, 4 of them present a female voice for the AD. In these cases, male characters are more numerous so this might explain the choice of female voice for the AD. Contrarily, only *Slumdog Millionaire* and *Borat* have a male voice for the AD, even though in both films there is a highest number of male main characters

Munich has only one female voice. In this film, parts in a secondary language are less frequent. In this case, the AD is delivered by a female voice, who is also delivering all the audio subtitles, for both female and male characters.

2.3. Audibility

A more technical feature, which is less related to the way audio subtitles are recorded or delivered by voice talents is the sound volumes of the AD/AST track in relation to the film soundtrack. In some of them, the audibility of the original line in a secondary language is almost lost even if the original is not completely removed.

In the film *The Science of Sleep*, the audio of the film soundtrack cannot be heard when an audio subtitled is delivered. This effect causes a noticeable contrast when the audio subtitle is finished and the original line is still on, letting the original line be heard at a very high volume in comparison to the volume it reached along the duration of the audio subtitle. This effect can sometimes surprise the user, who can notice the changes in the volumes.

In the film *Slumdog Millionaire*, the originals are barely heard. They are not completely removed but the AST does cover most of the original dialogues completely. In an argument between the two boy brothers (Scene 7. 13:38-14:05), the original dialogues in Hindi are emotionally loaded but the volume is so low they almost disappeared. The voice in the AST is more flat, closer to reading-aloud characteristics. In this case, the kernel of the scene is lost for users using the AD.

More interesting is the case of the film *Borat*, where the voice-over effect seems to match the fictional documentary tone of the film, that might have been a purposeful request for the AST. The humorous tone of the film itself, the performed voice of the audio describer and the accentuated prosody in the AST relate perfectly to the topic of the film (a documentary about North-American culture for Kazakhstani audiences). Thus, AST are more than merely a service used to make a film accessible. In the film *Borat*, the AST could be considered to be part of the whole artistic unit.

In the rest of films, *Munich*, *Syriana* and *Everything is Illuminated*, the audibility allows the original line to be heard under the AST. The fact that the dialogues are more paced and structured allows as well for a better correspondence between the original line and its audio subtitle.

2.3. Isochrony

Overall, there is a common tendency in the six films. In most of cases, AST is not presented in synchrony with the original utterances. This allows the user to listen to the original voice of the actor after or before the audio subtitle is finished. This type of delay is also used as way of providing more room for other AD contents.

The following graphs include information on the percentage of audio subtitles for each of the types of isochronies proposed: later, after, before, synchrony initial, final. A graph is provided for each of the film, which is later followed by a discussion on the types of isochronies predominant in the film.

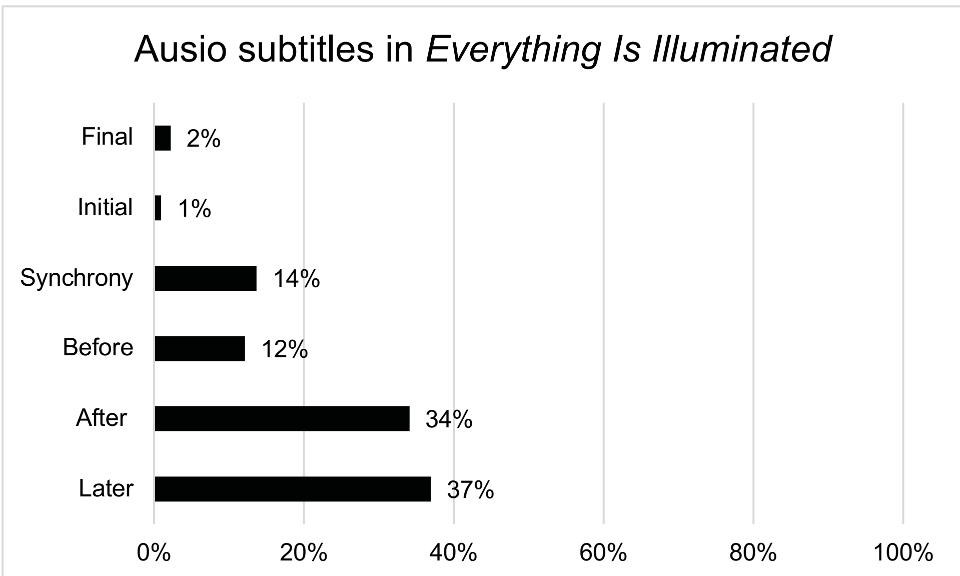


Figure 2. Percentage of audio subtitles for type of isochrony in *Everything Is Illuminated*

The film with *Everything Is Illuminated* is the film with the most audio subtitled minutes in the corpus. As shown in Figure 2, there is a tendency towards the delivery of the audio subtitles some time later in relation to their corresponding original utterance (37%). Very similar is the percentage of audio subtitles being delivered after the original utterance has been completely said (34%). The other strategies followed in this film are placing the audio subtitles in synchrony with the original (14%), and before the original utterance (12%).

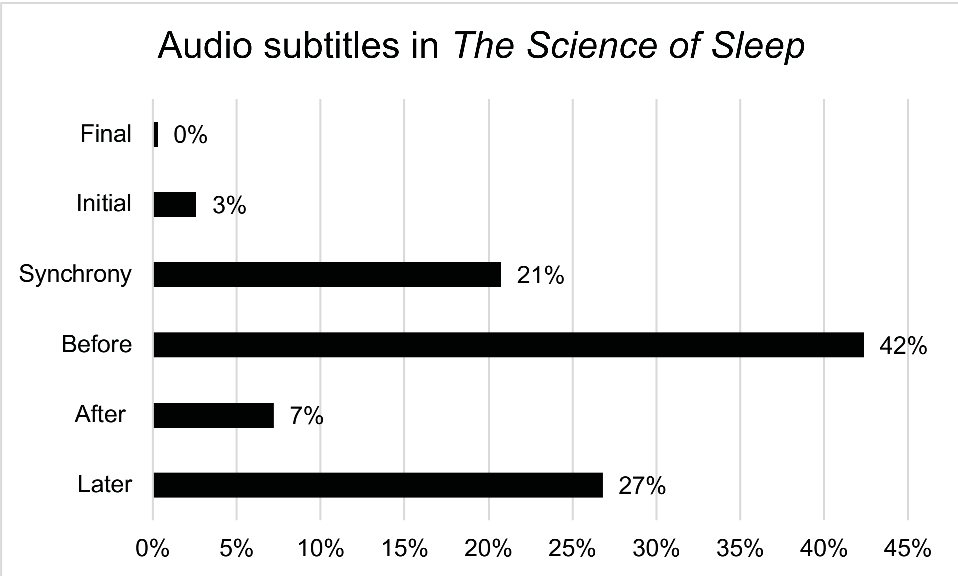


Figure 3. Percentage of audio subtitles for each type of isochrony in *The Science of Sleep*

On the other hand, in the film *The Science of Sleep*, the majority of audio subtitles are delivered before they are actually heard in the original language (42%) (see Figure 3). This percentage is followed by those audio subtitles being delivered some time later than the original utterance can be heard (27%) and those in perfect synchrony (21%). The rarest isochrony types in this film are after (7%) and initial (3%). As in the previous film, there is a trend observed followed by the audio describer and/or audio subtitler of this film, since almost half of the audio subtitles have been delivered before the original line.

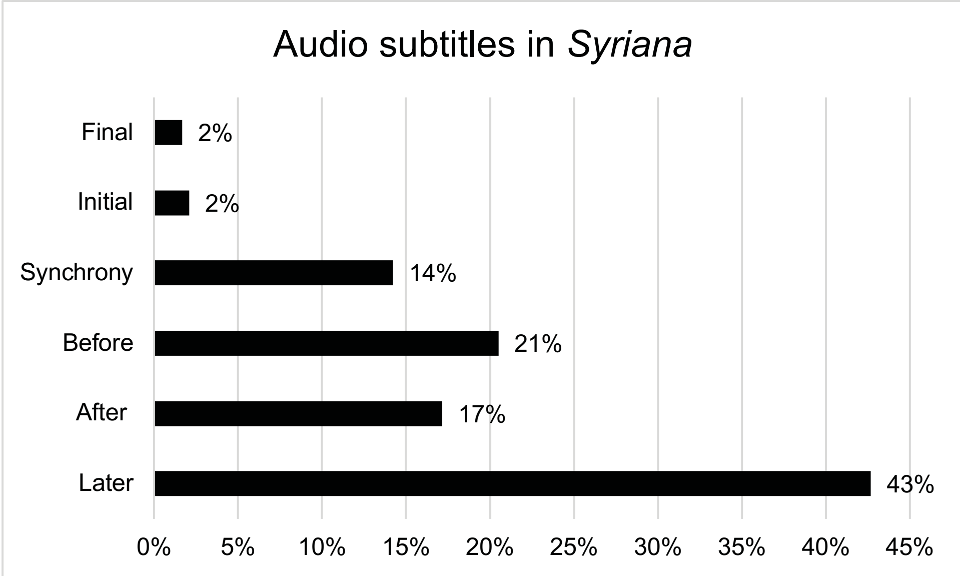


Figure 4. Percentage of audio subtitles for each type of isochrony in *Syriana*

In *Syriana* (see Figure 4), most of the audio subtitles are delivered later in comparison to the corresponding original utterance (43%), followed by those audio subtitles delivered before the original utterance (21%), those coming one the original is finished (17%) and those delivered in total synchrony with the original utterance (14%). It could be stated that a trend can be observed and that there has been an intention of providing AST in a certain way.

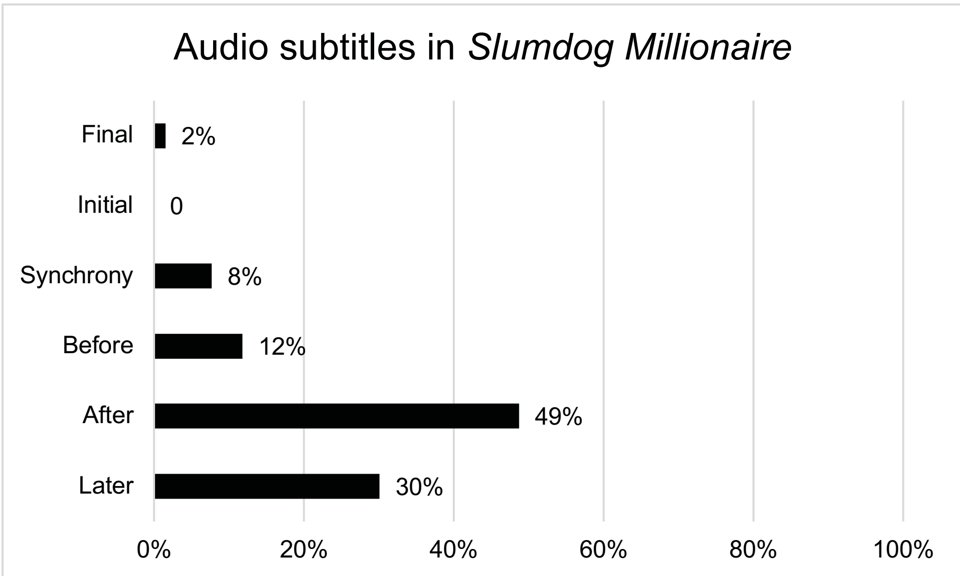


Figure 5. Percentage of audio subtitles for each type of isochrony in *Slumdog Millionaire*

In all, in the film *Slumdog Millionaire* (see Figure 5) almost half of the audio subtitles are delivered after the original utterances, when these are not heard any more (49%). The rest come some time later than the original (30%), before the original (12%) or in synchrony (2%). Again, there is a trend recognisable in the approach taken by the professionals in charge of the AD and AST content.

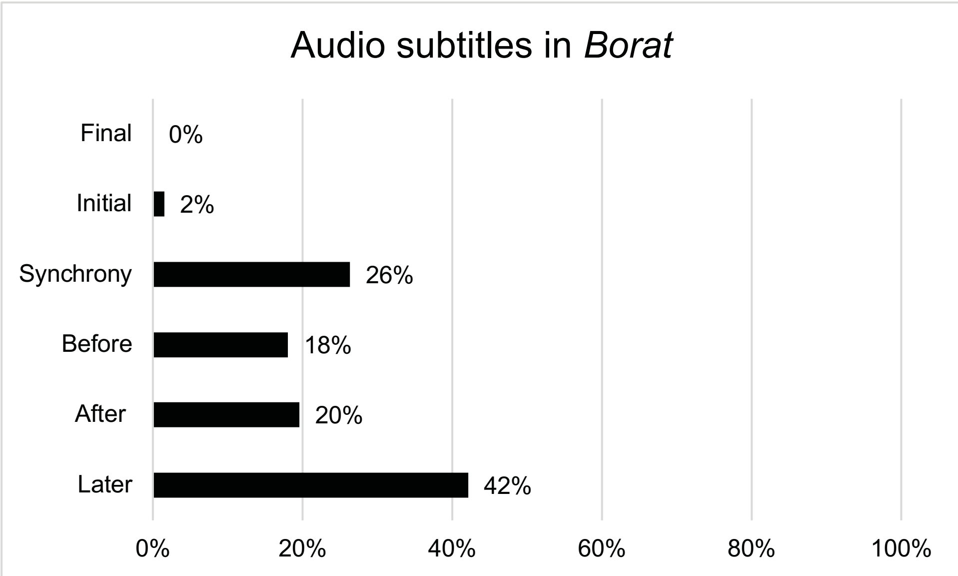


Figure 6. Percentage of audio subtitles for each type of isochrony in *Borat*

In the film *Borat* (see Figure 6), audio subtitles are normally delivered sometime after the beginning of the original utterances (42%). The rest of audio subtitles are delivered in synchrony with the original utterance (23%), after the whole original utterance (20%) or are introduced before the original (18%).

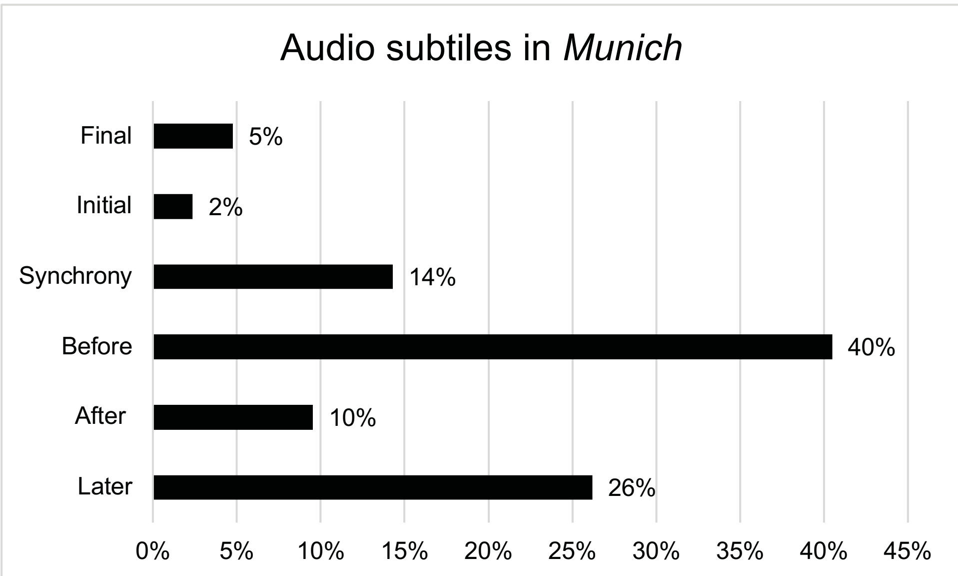


Figure 7. Percentage of audio subtitles for each type of isochrony in *Munich*

The same pattern is present in the scenes with AST in *Munich* (see Figure 7). Most part of the audio subtitles are presented before the beginning of the original utterance (40%), followed by those subtitles being displayed with some delay in relation to their corresponding original utterance (26%) and those played in synchrony (14%) or appearing after the original (10%). Again, the number of audio subtitles presented before suggest that this was the main strategy of the audio describer or audio subtitler.

2.4. Prosody

In general all the AST service within each one of the films seems to be consistent and has the same features when it comes to prosody. That means that the way audio subtitles are delivered show a clear trend or style in all of the films. This said, 4 out of the 6 films in this corpus have shown audio subtitles with prosodic features closer to acting, in the sense that they try to convey the interpretation in the original lines through a fake orality closer to dubbing. The next paragraph presents an analysis for each of the films.

In *Everything is Illuminated*, both Alex, who is presented as a young adult wearing gold necklaces and tracksuit in a style which could be defined with the derogatory adjective “chav*”* and his grandfather, who is a countryside man, are provided with audio subtitles that show prosodic features that match their social background in comparison to Jonathan’s English, who always wears a black suit and glasses and has perfect pronunciation. Apart from the prosody used in the AST voices, which stands as a recreation of the original tone, the accent used in the English AST is from RP pronunciation. In Alex, the matching of the prosody in the AST is in better correspondence, since he also speaks English with a scally or chav tone, which is respected in the AST, yet without the Ukrainian accent.

Borat is a very good example of a film where prosody both in the AD and the AST plays an important role in the whole narrative of the contents. For the amalgam of absurd and hilarious events, the voices used in the delivery of the AD and the AST present a humorous tone. The AST could be described as exaggeratedly acted but as a matter of fact, they are a reflection of the performances of the actors in the film, particularly those of Borat and Azamat.

In the film *The Science of Sleep*, even if some French-speaking characters being audio subtitled in English have contributed in English with a strong French accent, accents are not imitated. However, some of the prosodic features are recreated, a general example could be Scene 28 (1:12:50-1:13:33) in *The Science of Sleep* where there is some whispering which is also represented in the AST.

In *Syriana*, some acting is found in the prosodic features of the delivery of AST. First of all, the film falls far from having a humorous tone and secondly, in most of the scenes there is not a multilingual context, i.e. the characters are maintaining conversations in the language shared by the rest. When the emir is talking to his guests in a party he uses a very solemn tone which is portrayed in the English AST, paced and grandiloquent (Scene 7. 21:37-22:06).

The same happens in *Slumdog Millionaire*, which presents quite tragic scenes and situations. All characters in the flashback (when they are children) speak in Hindi. When they are adults, they speak English and they have a very strong Indian accent. The AST provided for the children speaking in Hindi does not have any trace of the Indian accent.

In the film *Munich*, AST are rarer in relation to other films. In this case, the prosody is very flat, and subtitles are just read aloud in a style more related to a documentary or to reading aloud. AST does not try to mirror the prosody of the original.

2.5. Assignation

For assignation three of the films here analysed use a reported assignation in which the AD content points who is talking. The other three introduce audio subtitles without further information, in what is labelled direct assignation.

In *Borat*, where the voice used in the AD and the AST of the two protagonists is the same, the AD introduces the AST with a "*she/he says*" strategy. In Scene 6 (5:52-5:57) the utterance by a secondary character, the AD goes "*Borat's neighbour says:*". This neighbour is also a male character, thus, he is matched with a male voice which seems to be the same one of the AD. Another example in the same film (Scene 8, 10:12-10:49) is when the two main characters, Borat and Azamat, are in a bar and Azamat is telling Borat off. The audio describer introduces the AST saying: "*He is berated by Azamat:*" and then the AST is delivered by the same voice with a change in the prosody.

Slumdog millionaire presents, as well, some characters with similar voices (the children) audio subtitled by the same male voice. Furthermore, some of the scenes in which the children are portrayed are fast changing and dialogues are fast. Due to such particularities, the AD is a key tool when recognising characters by mentioning who the character talking in that moment is before the audio subtitle is provided. A clear example is Scene 1 (5:53-6:30), where all children are being followed by policemen, and they all talk and shout. The AD track is mentioning the name of the character speaking: *Jamal says…, Celin says…*

The only film using only one voice for both the AD and AST content is *Munich*. Furthermore, there is little change in the prosodic features of the audio subtitles. Therefore, each of the audio subtitles is introduced by the AD with the word *subtitles*. One of the examples is found in Scene 2 (28:26-29:39), where Yvonne is talking to Avner and Andreas. The AD mentions the structure of the scene, the where and who, and then introduces Ivonne's words with *subtitles*.

In the films *Everything Is Illuminated* and *The Science of Sleep* there is no prior introduction of the audio subtitles from the audio describer. Audio subtitles are just delivered in what it is considered to be a direct style.

In *Syriana*, this is also the practice, but there is one exception in the film. In one of the scenes (Scene 17, 1:04:49-1:04:47), one of the main characters, Bob, is being tortured. Although the native language of the torturer is Arabic, the character speaks in English with Bob, who is American, and only changes to Arabic in some isolated interventions. The audio subtitle is introduced with a *he says*.

2.6. Textual Features

Subtitles, as a type of written text, are issued to be read. Therefore, subtitles include some translation strategies aimed at the reduction and omission of what is spoken, normally due to time and space constraints. One of the questions here posed when analysing the films was whether or not audio subtitles were different from their corresponding subtitles.

This may include the recovering of oral features in AST that may have been lost in the written subtitles or the expansion of the AST in relation to their written counterparts. The main trend observed is that, in general, audio subtitles are exactly the same as the written subtitles. Only some minor examples are found such as in *Slumdog Millionaire*, for example, in Scene 1 (5:53-6:30) written subtitles read “We won’t”, whereas the AST says “We don’t”. No important changes have been observed in this feature that should be reported. No importance of this feature has been observed for it to be reported.

3. Conclusions

In the presentation of the results, some main trends have been identified and the descriptive analysis has been extremely useful when understanding AST in general and identifying the strategies that are here exposed here. In the next section, the main trends identified for each of the aspects here analysed here will be exposed: voices, prosody, assignation, audibility, isochrony, and textual features.

There is a tendency towards differentiating the AD, with the use of the dichotomy of female and male voices, from the rest of the characters of the film. This strategy is found in *Syriana*, *Everything is illuminated* and *Munich*, where there are a lot of male characters in comparison to female (in *Syriana* only Prince Nasseer’s wife; in *Everything Is Illuminated*, Lista and the bartender; and Yvonne in *Munich*). However, not all films have been adapted following the same pattern, in *Slumdog Millionaire* and *Borat*, there are more male than female characters and they use a male AD voice. The case of *The Science of Sleep* shows a similar number of male and female characters. In this case the AD voice is female and the reason can be found in the main character of the film, who is Stéphane, who has a male voice.

One of the strategies found after the analysis consists of the use of two voices, one female and another male. This strategy has been followed in 4 films in opposition to 2 that do not follow this principle.

Regarding assignation, in three films of this corpus, it takes the form of a direct speech. AST are introduced without prior notice after the audio describer provides information about the structure of the whole scene. In the rest of the films the AD uses some type of indication that subtitles are going to be read by mentioning "*he/she says*", "name *says*" or stating that subtitles appear.

In the corpus proposed, the films with more audio subtitled minutes are those using a direct style (*Everything Is Illuminated*, *The Science of Sleep* and *Syriana*). This can lead to the suggestion that when AST is a big part of the audio described content, a direct style is preferred, maybe to avoid repetition of cues such as the aforementioned, or because it can provide more naturalness and fluency to the film.

As far as audibility is concerned, in five out of six films in the corpus, the original lines can be heard under the AST. That means that in general the common practice is that of a voice-over effect. However, in the film *The Science of Sleep* the original line is removed when an audio subtitle is delivered, and this also causes the original soundtrack to be lost, resembling to a dubbing effect.

Concerning isochrony, the common tendency in the six films is placing the beginning of the audio subtitles later than the beginning of the original utterance (here called *Later*). This is the case of three films in the corpus (*Everything Is Illuminated*, *Syriana* and *Borat*). The rest of the films introduce the majority of audio subtitles before the original utterance (*The Science of Sleep* and *Munich*) or after the original has finished (*Slumdog Millionaire*).

The type of isochrony with the highest percentage establishes the main trend in the film for AST delivery. However, as seen in the graphs showing the percentages of all the films, there is no film in which all audio subtitles have been delivered with the same type of isochrony. This varies according to the construction of the scene and the amount of aural information to be conveyed, namely, written subtitles to be translated, the AD contents, and the soundtrack.

It is worth stating that delivering audio subtitles in perfect synchrony with the original line is not the main trend in any of the films, so there is a tendency in the corpus to deliberately show traces of the original dialogue, thus pointing at a voice-over effect instead of a dubbing effect in terms of isochrony.

Another item that has been analysed is prosody***:*** four of the six films analysed in this corpus use an AST which include prosodic features that differ from those used when a text is read aloud. Therefore, there has been an intention to mimic prosodic features aimed to portray emotional load or intention. The degree of acting of the audio subtitled dialogue tries to match the film tone. Humorous films, such as *Borat*, or containing funny dialogues, such as *Everything is Illuminated*, tend to present AST with a lot of performance. In the film *Syriana*, much more dramatic, dialogues are delivered according to the acting in the original lines. Finally, in *The Science of Sleep*, some acting has also been identified, the humorous tone in some of the scenes has been matched. An example is found in the interventions of Stéphane's colleagues: Guy, Serge and Martine.

AST can also contribute to the artistic experience of consuming an audiovisual product instead of only standing as a mere translation aid. Examples such as the film *Borat* show how using acting when delivering AST can create a better synergy between the narrative nature of the film and the accessibility services that fulfil the specific needs of some audiences.

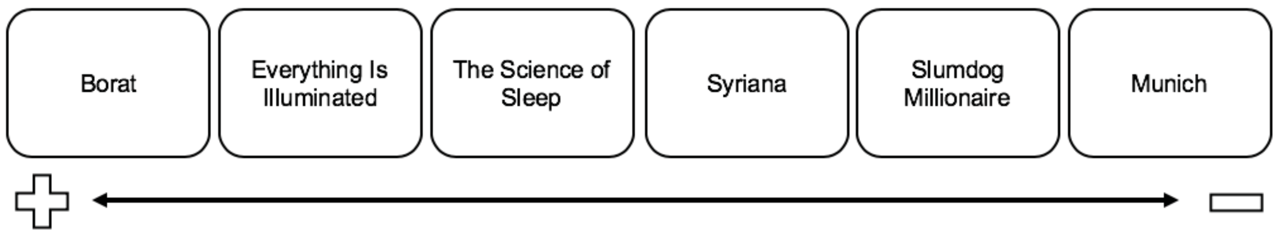


Figure 8. Films organised in terms of acting in the AST Service

In other cases, AST is delivered with less marked prosodic features, identified as read-aloud prosodic features within the framework of this analysis. It has to be taken into account that AST in this corpus is human-voiced, and the prosodic features cannot be divided between the acted/read-aloud duality. There are also some degrees in the acting of each of the AST services. Figure 8 shows the films of the corpus ordered according to the acting in the AST service.

Regarding textual features, in the corpus AST tend to be read exactly the same as the written subtitles. This feature is kept when written subtitles contain grammatical mistakes (Stéphane in *The Science of Sleep*) and idiolectic (Alex in *Everything Is Illuminated*) features of the characters. This contributes to intensifying the exoticism and foreignness of the scenes. Particularly for multilingual films, in which different cultures and nationalities are portrayed, this could be a part of the strategy to maintain the core of the film.

Linking the analysis with the theoretical categorisation proposed in Chapter 2, assumptions can be made about the main trend in the representation of multilingualism that the AST service of the films follows. The most popular representation of multilingualism through AST is “selective reproduction” without further information on the audio describer’s behalf. This allows the audience to understand that there’s a translation. In the case of *Borat*, the representation of multilingualism goes for a “verbal transposition” in which there is a will to imitate and even reproduce odd grammatical structures that characterise the idiolects of the characters. The film *The Science of Sleep*, uses a form of representation of multilingualism that falls between “selective reproduction” and “explicit attribution”, in which a dubbing effect is used and the original language has been mentioned in the beginning. In films such as *Syriana*, in which more than one secondary languages is spoken, “selective reproduction + language information” is used to highlight if the character is speaking in Arabic, Farsi or Urdu.

Chapter 6

Study 2: Measuring User Experience

Chapter 6
Study 2: Measuring User Experience

The past two chapters focused on the descriptive analysis of audio subtitled films. Chapters 6 and 7 focus on the experimental part of this thesis. This chapter presents the methodology used and the next one discusses the results.

The main aim of the experiments in Study 2 is the comparison of the emotional activation provoked by the two AST voicing strategies in a blind and partially sighted audience and also the preferences of this audience for one of the voicing strategies. Furthermore, it aims to test the differences in emotional activation triggered by AST in blind and partially sighted participants and written subtitles in a sighted audience.

The Study 2 is divided into two phases. Phase 1 was performed to select the stimuli and validate them by means of an online validation. Phase 2, in which the validated stimuli were used, is composed of Experiment 1 and 2. Experiment 1 was performed with blind and partially sighted participants and audio subtitled stimuli. Emotional activation was measured as well as preferences. Experiment 2 was conducted with sighted participants and subtitled stimuli. In this case, only emotional arousal was measured.

The methodology is presented chronologically. **Section 1** explains Phase 1, which had the purpose of identifying the stimuli for the following steps of the process. In **Section 2**, Phase 2 is explored**.** In this phase the identified stimuli were used to test the impact of the aforementioned effects on participants. Phase 2**,** in turn, is divided in two parts: Experiment 1 and Experiment 2. Experiment 1 was performed with blind and partially sighted participants with audio subtitled clips, whereas Experiment 2 was performed with sighted participants and the same subtitled clips. Both experiments share the same experimental protocol. In the following paragraphs, the two phases of the experiment and the two parts (Experiment 1 and 2) within Phase 2 are described in detail. At the end of Section 2, a diagram (Figure 22) is provided on the development of Phase 1 and Phase 2. Finally, **Section 3** provides a report on the approach taken to the analyse the raw data obtained.

1. Phase 1: Online Validation

The selection of audiovisual stimuli to be used in the Phase 2 was performed by means of an emotional validation. Participants were asked to complete an online survey remotely. Eight clips were uploaded online for validation and divided in two separate surveys that allowed each of the surveys to have four clips and to be responded separately, thus reducing the duration of the tests. Participants were presented with one of them and were afterwards encouraged to respond the other. This section summarises the main features of the online validation regarding: participants, materials, stimuli and duration.

1.1. Participants

Since this phase was performed remotely online, the objective was obtaining as many participants as possible with the initial objective of reaching 100 participants. Eventually, the number was overtaken. One of the surveys reached 110 participants while the other one 121 respondents. No demographic data were controlled in order to reach a wider sample.

1.2. Materials

Information about the materials used is divided between apparatus, referring to all technological devices, and instruments, describing the surveys used for the validation of the clips.

1.2.1. Apparatus

The survey was created with Survey Monkey. This software allowed for the creation of two different surveys with links between each other. Participants accessed them through a device connected to the Internet. Survey Monkey allowed for the restriction of access to the survey by means of IP monitoring. Each of the surveys could only be done once with the same IP, this way the probability of malfunction or inaccuracy greatly decreased. This software also generated graphs that helped in the identification of the most suitable stimuli to be used in the Experiments 1 and 2 of the following Phase 2.

1.2.2. Stimuli

Scenes from the Polish TV series *Wojenne Dziewczyny* [War Girls][7] (2017), broadcast by the Polish national channel *Telewizja Polska 1* (TVP1) were used in all phases of the experiments. Polish was selected because one of the selection criteria was to use scenes in a language unknown to the audience. The objective was to reproduce a scene in a multilingual content where secondary languages are subtitled, and Polish is normally unknown to Spanish audiences.

Extracting all the stimuli from the same audiovisual product has a series of advantages. First, it provides cohesion in terms of acting, character voices, plot, and musical and ambience effects. Second, the project NEA, in which this thesis is developed, reached an agreement with the broadcast company by which the scenes could be used and treated (with subtitles, AD and AST) for research purposes, solving copyright issues.

The Series

The series *Wojenne Dziewczyny* (2017) (translated for the purposes of this thesis as *Mujeres en guerra*), tells the story of Irka, Marysa and Ewa, three young women who get involved in the resistance forces against the Nazi occupation in 1940s Warsaw. Irka is a chemistry teacher, Marysa is a Jewish Pianist and Ewa is a former thief from the outskirts of the city. Despite their differences, they join their skills and efforts to fight against the occupation. The series was directed by Michał Rogalski, and was released in 2017 by the Polish cannel TVP1. The series is currently in the second season, aired in 2018.

1.2.2.1. Identification of the Scenes

The scenes were selected after watching the whole first season of the series on the website of the broadcasting channel TVP1[8]. The selection was based on two main requirements:

1. Each scene should have a duration longer than 2 minutes.
2. Each scene should be self-contained.

[7] https://vod.tvp.pl/website/wojenne-dziewczyny,28767487
[8] https://vod.tvp.pl/website/wojenne-dziewczyny,28767487

The emotional validation had the objective of confirming that the emotion identified by the researcher in each selected scene is commonly accepted among audiences.

The scenes were taken from episodes 2, 3, 5, 6 and 12 of the first season and show an emotion or the absence of it: anger, fear, sadness, none, or as it is referred to in this thesis, neutral (although not an emotion, it portrays its absence). As already explained in Chapter 3, the reason behind the selection of the emotions is based on previous experiences of emotional activation and AV contents.

The scenes selected were each around 3 minutes long and are self-contained, so that participants can follow it easily. In the following section the scenes identified for the validation are explained in detail and are accompanied by the emotion that was associated to them, the episode from which they were extracted, and their exact duration.

1.2.2.2. Treatment of the Scenes

After selecting the scenes, the broadcaster was contacted. They provided the scenes from the archive with a water mark and gave permission to use them and manipulate them for research purposes. A professional translator was hired to provide subtitles into Spanish.

1.2.3. The Scenes

Scene 1: *Sadness (Episode 2. 2.41")*

Witold explains to his mother, her sister Irka, Marysa and Ewa that the Nazi have captured two of their colleagues. Ewa goes dramatically to her room, Irka follows her to see what happened.

Scene 2: *Sadness (Episode 5. 2.22")*

Ewa is examining some bullets that are hidden in a package full of sand. Witold, head of the resistance organisation and for whom she is developing some feelings, comes in after the landlady announces he is there to see Ewa, who has recently lost her ex-boyfriend in Nazi's hands. They have to pretend that they are engaged so as to not raise suspicions.

Scene 3: *Fear (Episode 6. 3.27")*

Irka is trying to smuggle an explosive substance and mistakenly ends up at the apartment of one of her ex-professor, who has changed his surname to a German one and is a supporter of the Nazi regime.

Scene 4: *Fear (Episode 12. 3.16")*

Irka's mother and Ewa are being interrogated by a Nazi general.

Scene 5: *Anger (Episode 5. 3.04")*

The three girls, Ewa, Irka and Marysia go to the apartment of a Polish girl they know. She is having an affair with the Nazi general who killed Ewa's mother. Ewa wishes to take her revenge and they wait for him to come.

Scene 6: *Anger (Episode 12. 2.14")*

Ewa and Irka interrogate a man, gun in hand, because they want their money back. He is doesn't want to pay them back and is hiding more things not from them.

Scene 7: *Neutral (Episode 3. 2.11")*

Irka is home, studying for an exam. Marysia and Ewa come home from a walk. They speak about getting a new gun for Ewa by flirting with a German soldier.

Scene 8: *Neutral (Episode 3. 2.11")*

Marysia is playing her viola in her room. Irka comes in and they talk about Ewa being in love with Witold, Irka's brother.

In Table 17 the information about each of the scenes is found. Each scene is followed by the emotion intended, the episode from which it was extracted and finally, the duration in minutes and seconds.

Scene	Emotion	Episode	Duration
1	Sadness	5	2.41”
2	Sadness	6	2.22”
3	Fear	6	3.27”
4	Fear	12	3.16”
5	Anger	5	3.04”
6	Anger	12	2.14”
7	Neutral	3	2.11”
8	Neutral	3	2.11”

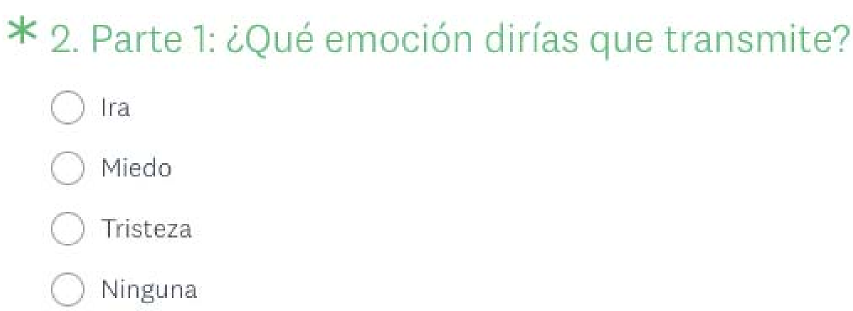
Table 17. Summary of the scenes identified for the online validation

1.2.3. Instruments

The survey used for the validation contained two parts:

1. A question concerning the emotion experienced during the viewing of the scene. A closed list was provided: fear, anger, sadness and neutral emotion. *¿Qué emoción dirías que has experimentado?* [What emotion has the clip provoked in you?]. A list of 4 emotions was provided: *ira* [anger], *miedo* [fear], *tristeza* [sadness] and *ninguna* [none].

2. The first two parts of the SAM questionnaire (Bradley & Lang, 1994), valence (Figure 9) and arousal (Figure 10). This self-report instrument has been introduced in Chapter 3 as a questionnaire concerned with the emotional status of the participant.


* 2. Parte 1: ¿Qué emoción dirías que transmite?

- Ira
- Miedo
- Tristeza
- Ninguna

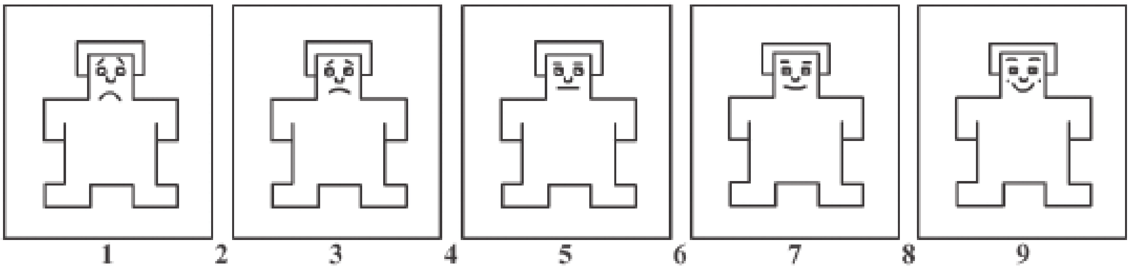


Figure 9. Valence dimension of the SAM questionnaire

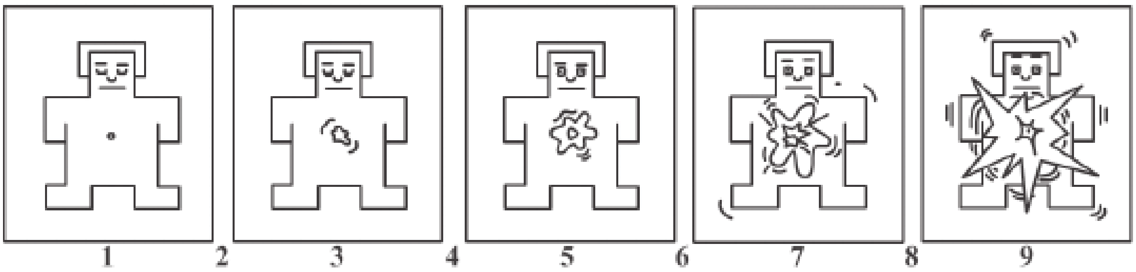


Figure 10. Arousal dimension in the SAM questionnaire

1.3. Duration

Each of the surveys consisted of four clips of around 3 minutes long, and then a maximum of thirty seconds to complete the questions for each clip. According to a pilot test with 4 participants the expected time of the completion of one survey was 14 minutes and the completion of the surveys forms was around 28 minutes.

1.4. Ethical Permission

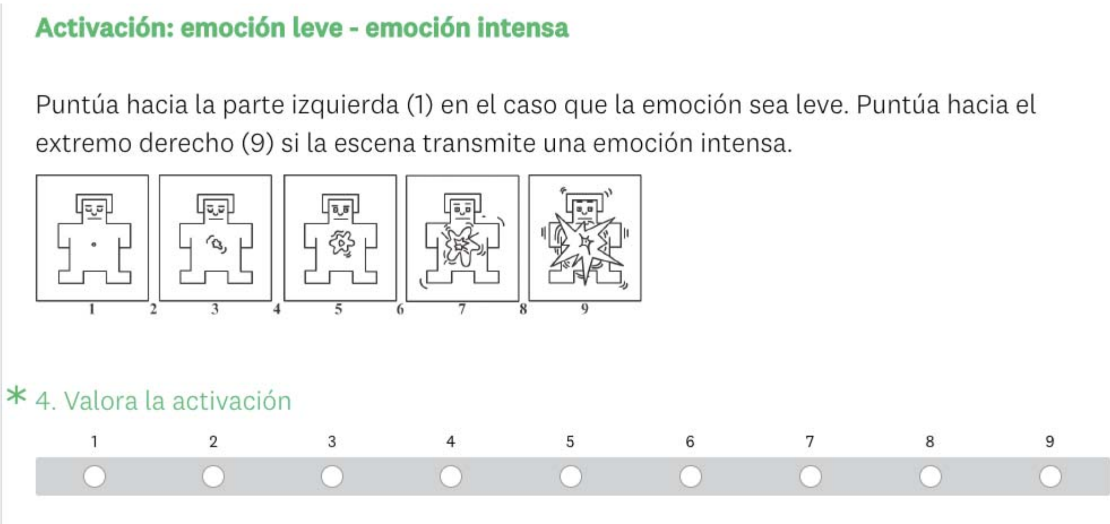
The validation was approved by the ethical committee of the Universitat Autònoma de Barcelona. The participants, before starting validating the clips, had to read an information sheet with the description and aim of the experiment and express their explicit consent (see Section 2 in the Annexes).

1.5. Contacting the Respondents

The two surveys were advertised by means of social networks and direct contact with participants. The survey included a password that was provided at the beginning of the surveys as captcha. At the same time, the software traced the IP of the devices so the surveys could only be accessed once from every device.

1.6. The Tasks

Before responding to the survey, the participants were presented with the ethical consent that had to be signed. Then the tasks were repeated after the visualisation of each of the clips (4 in total). The participant watched the clip and then were presented with the question about the emotion experienced [Rate towards the left (1) in case the emotion is weak. Rate towards the right (9) if the scene transmits an intense emotion]. After that, they had to respond to the two scales corresponding to the first two dimensions of the SAM questionnaire (Figures 9 and 10).



After the four clips were visualised and their corresponding tasks were finished. The participant was presented with the link to the other survey with another 4 clips. One or the other survey were sent randomly as first survey to the participants, so the replies were balanced.

1.7. Pilot Validation

Once all the online validation had been designed and set up in Survey Monkey. Four participants were contacted to complete it and provide feedback on the development of the survey. Overall, the survey was correctly understood and only small changes had to be made in relation to the terminology, that had to be coherent along the whole survey.

Different verbs were used to refer to the emotion “experienced” by the participants in the first question of the survey, such as *sentir* [feel]. The same was found in the SAM questionnaire, in which the final verb used is *transmitir* [transmit], but verbs such as *sentir* [feel], *experimentar* [experience] or *suscitar* [provoke] were used as synonyms.

1.8. Final Clips

Out of the two scenes for each emotion (sadness, fear, anger, and neutral), one was chosen: the scene which acquired higher percentage of respondents identifying the emotion of the clip and, secondly, which was more representative in valence and arousal according to the literature. Among anger and fear, two emotions that might be confused when analysing their psychophysiological effects (see Chapter 3), the one with higher scores was chosen.

In Table 18 the scenes selected are shown together with the following information: the percentage of responses coinciding on the emotion intended, the average score of valence for that scene, and the average score for arousal. The clip corresponding to “neutral” proved to achieve a medium valence and a very low arousal. Fear was the emotion with lower score in valence, thus implying that the emotion was experienced as negative, and the arousal proved to be quite high. Sadness also was related to a low valence score and a higher arousal than neutral but lower than fear. In general, the results corresponded with the information provided in the theoretical framework (see Chapter 3).

Scene	Emotion	Percentage	Valence	Arousal
7. Episode 3 | 2.11”	Neutral	81.82%	5.58	3.73
3. Episode 6 | 3.27”	Fear	71.59%	3.11	6.58
2. Episode 5 | 2.22”	Sadness	88.54%	3.14	5.42

Table 18. Online validated scenes and their ratings

2. Phase 2: The Experiment

Phase 1 obtained the stimuli for Phase 2, which was divided in two experiments: Experiment 1 was carried out with blind and partially sighted participants and audio subtitled clips and Experiment 2 was carried out with sighted participants.

2.1. Experiment 1

Experiment 1 used the stimuli obtained from the validation treated two different AST voicing strategies: voice-over effect and dubbing effect. This experiment was conducted with blind and partially sighted participants. In the next paragraphs its duration, participants and materials are explained.

2.1.1. Duration

Experiment 1 had a duration of around 60 minutes. The presentation of stimuli and relaxation times, which were automatized by means of the software controlling the experimental procedure, had a duration of circa 30 minutes. The other parts of the experiment, such as the welcoming of the participant, the explanation of the experiment, the demographic test and pre-test and final preference questionnaire, were more variable in terms of duration.

2.1.2. Participants

Experiment 1 was performed with 42 participants: 13 blind and 29 partially sighted persons (17 females and 25 males) aged 19-73 years (median=38). The essential recruiting condition was that partially sighted participants cannot access written subtitles and that they should rely on the AST to understand the message. Variety in the sight condition aims to portray the reality within the blind and visually impaired community.

Participants were reached through the Associació Discapacitat Visual Catalunya – B1, B2, B3 (Association for Visual Disability Catalonia - B1, B2, B3). This association provided a list of members from 18 to 50 years old who were contacted by phone. The experiment was explained and if they were interested, they were given an appointment to come and take part in the experiment. The association provided a space in their facilities to perform the experiments. People who were not members of the association could visit the facilities in order to take part in the experiment. This provided some freedom to contact other people outside the association by means of personal contacts.

2.1.3. Materials

Materials are divided between apparatus, that corresponds to all technological devices and software used for the collection of data and instruments, which refers to the pre-experiment forms and the questionnaires provided to the instruments to collect self-report data.

2.1.3.1. Apparatus

Data were collected by means of a psychophysiological data recording machine, Captiv L7000 from the French manufacturer TEA**.** The recording software was the one accompanying the machine. This device consists of a router (T-DAC, see Figure 11) which is connected with a cable to the computer and receives the data sent from two wireless receptors: one for the electrodermal activity (Figure 12) and one for the heart rate (Figure 13).

Figure 11. T-DAC, receiver device of the Captiv L7000

Figure 12. GSR sensor of the Captiv L7000

Figure 13. HR sensor of the Captiv L7000

On the other hand, the Tobii device (Figure 14) and corresponding Tobii software were used as procedure controlling software. This software monitored the presentation of stimuli, the instructions of the experiment, and the relaxation times. It was linked with both the Captiv L7000 devices and the Captiv L7000 software.

Figure 14. Tobii eye-tracker receiver and sensor

Both devices are two external boxes that are connected to a computer. The computer used for the experiment was a Toshiba Portégé 730-B-119 (Figure 15). The T-DAC from Captiv L7000, receiver of the signal sent by the two sensors, was connected only to the computer while the Tobii X6-20 was connected to the computer and to a sensor that records eye-tracking. This physiological input was not recorded in our experiments but the machine needed a former and symbolic calibration to start the protocol in its software.

Figure 15. Toshiba Portégé 730-B-119

The computer was connected by means of a VGA connector to a monitor, where the experimental procedure was played. The monitor used was a 23-inch HP Elitedisplay E232 (Figure 16), with a resolution of 1920 x 1080 at 60 Hz. The monitor had a socket from where the over-ear headphones were connected. The noise cancelling headphones used were the Sony MDR-ZX110NA (Figure 17).

Figure 16. HP Elitedisplay E232 monitor

Figure 17. Sony MDR-ZX110NA head-phones

A second computer was used to register all the information concerning the questionnaires completed by the participants. The second computer, a laptop MacBook Air, was prepared with the Google Forms sheets (Figure 18) corresponding to the different questionnaires. This computer had to be connected to the internet in order to effectively register all the answers.


Cuestiones previas

*Required

Cuestiones previas

1. ¿Género? *

Masculino

Femenino

Prefiero no contestar

Other:

2. Edad *

Your answer

Figure 18. Example of the Google Forms sheet

2.1.3.2. Instruments

Data were also collected by means of self-report questionnaires. In the following lines the instruments administered to the participants in Experiment 1 are explained and organised chronologically as they were administered.

Information Sheet

An information sheet of the experiment (see Section 2 in the Annexes) explaining the project and the experiment to the participant.

Consent Form

This consent form was approved by the ethical committee of the Universitat Autònoma de Barcelona (see Section 2 in the Annexes). The consent form stated that any personal details or any other data from the participants will be confidential and that the participant could stop the experiment at any moment without consequences.

Pre-questionnaire

The pre-questionnaire included:

a) Demographic questions
b) Questions about anxiety and caffeine, drug and medicament consumption (see Section 2 in the Annexes).
c) the Toronto Alexithymia Scale (TAS-20)(Taylor et al., 1985) to test their ability to identify and describe emotions in the self. Different translations of the TAS-20 questionnaire into Spanish are found. For the present study it was decided to use the validated Spanish translation in Martínez Sánchez (1996). The contents of the Alexithymia Scale can be found in the Annex of this thesis (see Section 2 in the Annexes).

T-SAM questionnaire

The T-SAM questionnaire was administered. This questionnaire is a haptic adaptation of the SAM questionnaire (Bradley & Lang, 1994) which was created within the framework of the present project (see, Iturregui-Gallardo & Méndez-Ulrich, in press) and is further explained in Section 2.2. in this chapter.

Final questionnaire

The final questionnaire was composed by questions on preferences. The questions centred on audio subtitling and its voices and, less importantly, on audio description. The original questionnaire was presented in Spanish, the translation of which can be found in the Annexes (see Section 2 in the Annexes).

2.1.3.3. Stimuli

Three different clips were used in the protocol of Experiment 1. The clips were extracted from the validation described in Section 1 in this chapter. In Experiment 1, such clips were audio described by a professional audio describer. Once the three scripts were received, the materials were sent to a dubbing studio (Escola Catalana de Doblatge (ECAD) [Catalan Dubbing School]) where the AD and AST were voiced by professionals. Each scene was audio subtitled in two different versions: one with a dubbing effect and another one with a voice-over effect (see Chapter 2 for further information on the two effects).

After the clips were modified in the dubbing studio, the stimuli were 6: sadness clip with a dubbing effect, sadness clip with a voice-over effect, fear clip with a dubbing effect, fear clip with a voice-over effect, neutral clip with a dubbing effect, and neutral clip with a voice-over effect as seen in Table 19. The stimuli were treated with only three voice talents: 1 female-voice for all female characters, 1 male voice for all male characters, and a third male voice for the AD

Scene	Emotion	Effect
2	Sadness	Dubbing
2	Sadness	Voice-over
3	Fear	Dubbing
3	Fear	Voice-over
7	Neutral	Dubbing
7	Neutral	Voice-over

Table 19. Experimental clips

Each clip was preceded by an aural explanation of the context of the scene. By providing an explanation, the participant could identify better with the action and the emotion of the scene.

2.1.4. Procedure

The presentation of the stimuli was done by preventing the participant from watching the same scene twice but making them watch the two conditions (dubbing and voice over effect). Since the numbers of emotions and the number of conditions were not even, a randomising strategy was created. Out of this strategy, as shown in Table 20, 6 separated combinations were obtained. Each of the combinations was randomised by using the experimental procedure software (Tobii). This randomised presentation of stimuli combines the emotion of the clip (sadness, fear or neutral) and the condition of dubbing (A) or voice-over (B) effect. Six different combinations were created. Since there are two conditions (dubbing and voice-over) for three emotions (fear/anger, sadness, neutral), each participant watched two of the scenes with the same condition. In Experiment 1, the 6 protocols were repeated 7 times, acquiring the number of 42 protocols corresponding to the 42 participants.

	Fear	**Sadness**	**Neutral**	
1	A	A	B	**x 5**
2	B	A	A	**x 5**
3	A	B	A	**x 5**
4	A	B	B	**x 5**
5	B	A	B	**x 5**
6	B	B	A	**x 5**

A=dubbing effect, B=voice-over effect

Table 20. Randomisation of the stimuli presentation

2.1.5. Experiment Outline

The experiment developed as follows:

1. Participant enters the room.
2. Participant is told about the nature of the experiment and is informed about the procedure, where the sensors are going to be place, and the instructions for the completion of the T-SAM questionnaire.

3. Participant is read the consent form, presented in a form with questions that the participant has to answer with yes/no responses.
4. The two sensors are connected.
5. The participant orally completes the demographic and general questions. The researcher registers the answers in Google Forms questionnaire (this allows for a quick completion of the task and an easier reading of the results).
6. The participant orally completes the TAS-20 questionnaire. The researcher registers the responses on a Google Forms questionnaire too.
7. The participant is then provided with headphones. From that moment, the presentation guides the participant by means of instructions appearing on screen and recorded orally by the researcher.
8. The participant is first invited to a relaxation period induced by mood music, which lasts 8 minutes. This relaxation period is aimed at unifying the participants' emotional state and to record the base line.
9. The first clip is played.
10. The participant completes the T-SAM.
11. The participant is invited to a relaxation period (3 minutes) before the next clip. This is repeated two more times.
12. The participant can remove the headphones and disconnect the sensors.
13. The participant is told that two different strategies were used to voice the subtitles and that the AD had the same characteristics in the three videos.
14. The questionnaire on preferences is administered.
15. The participant is thanked for their collaboration.

2.1.6. Pilot Experiment

The Protocol was tested with three participants. The protocol was understood correctly and the register of the psychophysiological measures worked. The results obtained in the pilot participants were also used for analysis, since no changes were made to the protocol and they were valid.

2.2. SAM Questionnaire for Blind and Partially Sighted Participants

The haptic version of the SAM questionnaire (Bradley & Lang, 1994) was validated by means of a focus group. The interest in adapting a self-report instrument for the validation of emotions was also triggered by an external need, not only the purpose of this experiment. According to research, visual impairment has a negative impact on psychological and emotional well-being (Barr, Hodge, Leeven, Bowen, & Knox, 2012) and normally psychometric tests used for the emotional assessment of sighted people are used in braille or orally for the blind and partially sighted (Méndez-Ulrich, Prats-Basset, Yagüe, & Sanz, 2016). The graphic-nature of the SAM questionnaire makes it quite unique for its duration and aimed simplicity. The whole creation and validation process of the haptic adaptation of the SAM questionnaire is found in Iturregui-Gallardo & Méndez-Ulrich (In press) developed within the framework of the present project.

The adaptation of graphic elements into their haptic representation poses different challenges, the two main ones being: (1) the understanding of the elements represented and (2) the materials to be used. The challenges were overcome by taking into account the following details:

1. The understanding of the elements:

Not a lot of literature is found on the haptic understating of graphic elements. Some institutions have published their own guidelines on the production of haptic graphic elements. Examples of this are the guidelines proposed for the American Printing House for the Blind Inc (1997), The Braille Authority of North America (2011) and the Tactile Book Advancement Group (2006). From the instructions provided in these documents, some key points were extracted to apply to the design of the T-SAM:

- Elimination: only relevant components are to be portrayed.
- Simplification: the main purpose is for the user to comprehend.
- Space (avoiding clutter): there should be space between elements.
- 2D: 3D elements should be reduced to 2D when possible.

- References based on experience: a part of the object/concept will provide more information than the whole representation of the object/concept.
- Consistency: the stylistic decisions must be maintained along the whole product.

2. The materials:

Printing for haptic purposes is expensive. Relief printing is normally obtained through the use of moulds against a variety of supports, such as paper or plastic. The creation of moulds has normally elevated costs. Other potential options are negative or positive relief with lasers on hard supports, such as methacrylate or wood or the use of different raisins to obtain relief. The selection of the material and technique should always take into account the readability on the users' behalf. According to The Braille Authority of North America (2011), for instance, a minimum of 1 mm relief must be acquired.

2.2.1. The Creation of the T-SAM questionnaire

Two different adaptations were created. The process of creation departed from the drafts issued by hand. These drafts (Tables 19 and 20) took into consideration the elimination and simplification mentioned previously. The body of the avatar in the first two dimensions was removed. The shape of the faces was left in one of the questionnaires to compare the acceptance in the focus group. The facial features posed as well a problem as well (eyes, eyebrows, nose and mouth of the avatars), again the removal of this feature was left for the validation in the focus group. The numbers were included in braille, occupying the same space graphic numbers had.

After issuing by hand the two drafts on paper, both designs were vectorised and digitalised. The programme used was Adobe Illustrator CS6. The format was saved in PNG, which allows the image to exist without a background, required for the future printing process. The material used was UV gloss. This material is cheap in comparison to other methods and is easily available in normal copy shops and printing houses. The printer applying the UV gloss, a type of raisin, basically applies layers of gloss. In the case of the questionnaires, the process was repeated 7 times to obtain the desired relief. The material of the base onto which apply the gloss was a foam board covered in plastic. It was not advisable to work with paper-based supports, since the gloss, when is hot, could be absorbed by the support, thus reducing or distorting the relief and shape of the design.

2.2.1.1. The Focus Group

The focus group with 5 blind (3) and partially sighted (2) participants (1 male) (aged 26, 27, 28, 47 and 51 years old, average = 35.8) was aimed to distinguish which of the two versions (Figure 19 and Figure 20) was more understandable and which features needed readdressing. There was not criteria in terms of age, sex or sight condition since the purpose of the T-SAM was to serve as an instrument for emotional validation for all persons regardless of their specific needs.

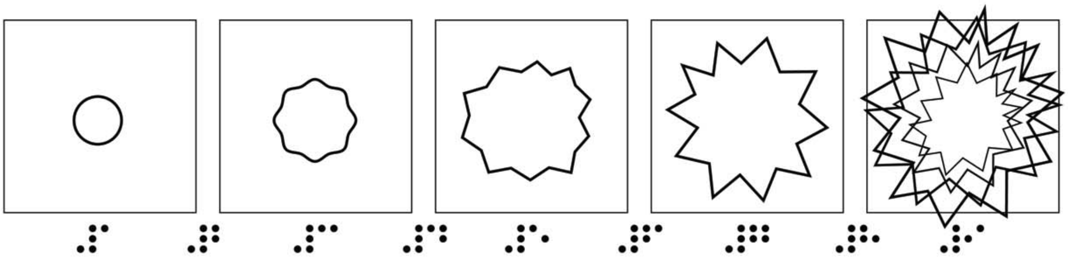


Figure 19. Version 1 of the T-SAM

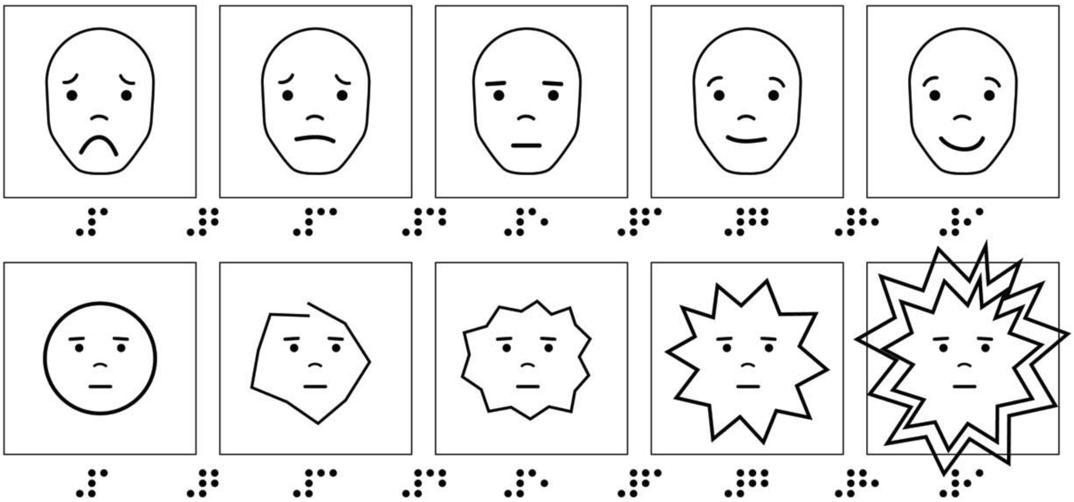


Figure 20. Version 2 of the T-SAM

Participants were informed that the focus group was going to be recorded and they all gave their consent. Afterwards, the moderator presented the two tactile versions of the questionnaires and allowed the participants to interact with the material in order to familiarise themselves with it.

After that, a series of questions (see Section 2 in the Annexes) were asked which were aimed at the assessment of, first, (1) the comprehension of the main objective of the questionnaire and of each one of the two subscales. Later the questions touched on particularly each of the scales and the tactile graphic items portrayed. For the first scale corresponding to the dimension of valence the questions concerned (2) the comprehension of the tactile representation of a simplified human face, (3) the difficulty to detect facial features related to emotional gestures (particularly lips and eyebrows), (4) the familiarity of each of the participants with the use of emoticons, since the face represented is an artificial representation of a face very similar to today's emoticons and finally, (5) the understanding of the shape of the face and the usefulness of tiny details as such as eyebrows, eyes and nose for understanding emotional gestures.

Secondly, questions were asked about the second series of images devoted to the dimension of arousal. The first question (1) focused on the understanding of the shapes and the faces found inside the shapes of one of the two versions of the T-SAM, later (2) the interest was placed on the relation between the shapes and the emotional activation translated into the ever-changing edges of the shapes. Since this representation of arousal goes from less to more, an open question about different ways of portraying such gradation (3) was asked.

Finally, the final section of questions embraced the whole questionnaire in general. The first question (1) asked the participants whether the questionnaires could be understood and answered only with the introductory explanation. The second (2) posed a question about which of the two versions was easier to follow and related directly with the third and last one (3) invited the participants to provide their opinion on which characteristics made the comprehension and completion of the questionnaire easier.

2.2.2.2. Results of the Focus Group

The results of the focus group are here presented according to the order in which the questions were posed to the participants.

Comprehension of the Questionnaire and its Subscales

Due to the specific needs of people who are blind or partially sighted, it was necessary to provide information not only about the functioning and meaning of each of the two subscales of the questionnaire (as is the case for sighted participants) but also about the physical characteristics of the material.

Participants understood the objective of the questionnaire and its two subscales.

First Dimension: Valence

Comprehension of the tactile representation of a simplified human face proved to be somewhat problematic. This first question related directly to the familiarity of participants with the meaning and form of emoticons. Whereas younger participants who had low vision or had been sighted in the past quickly recognised the faces and their gestures and their meaning was supported by their comprehension on the 1-9 scale, participants who were completely blind from birth had trouble understanding the faces. Two blind participants (aged 47 and 51) suggested that for them such images were no support and that they preferred to only have a braille numerical scale and an explanation from the researcher. However, since this questionnaire is designed for all kinds of user, whether blind or with low vision, having both options was considered the better solution. In fact, the T-SAM questionnaire is also suitable for sighted audiences since its simplification provided a quicker understanding of the tasks.

People with low vision also made some comments. Even if they could perceive the augmented and simplified lines of the faces represented, the relief helped them to understand the images, as well as the frame delimiting each face. They also preferred having the profile of the faces. As far as numbers in the scale are concerned, people with low vision suggested that they could benefit from the inclusion of augmented graphic numbers apart from braille (not all of them understand braille, or at least do not understand it with ease).

Even if blind participants, who were not familiar with emoticons, suggested that they could not understand such representations, they could understand faces and gestures with the guidance of the researcher. However, they made significant recommendations about the artificial representation of facial features.

The profile of the face helped understand the image**;** the nose was represented as a tiny horizontal curve and this is very artificial, blind participants understood the nose as a straight vertical line that separates a face in the middle; the tears were understood as a sign for negativity; the mouth, that was shaped as a single line that curved depending on the valence of the emotion brought some problems. Blind participants understood the mouth as two different lines corresponding to the two lips, not to the line formed by this two, which is normally associated to the mouth in emoticons.

In general, a proper explanation and some guidance around the tactile elements of the questionnaire were the key factor for the understanding of the whole questionnaire. All participants understood their task in this first series of images. It was confirmed that the community embracing blind and partially sighted people is very varied, and that some solutions and new options must be included in the T-SAM questionnaire to fulfil all the different needs.

Second Dimension: Arousal

Questions were asked in relation to the second dimension. They were related to abstract representation in the form of changing shapes. Once the meaning of the scale was explained to the participants, those with low vision clearly understood the relationship. Those accessing the subscale solely through touch, required a more focused explanation and some guidance which highlighted the changing feature: the edges of the shapes. Since the images found on this scale are a metaphor of emotional intensities that are based on an abstract concept, it was understood relatively quickly. As far as the faces inside the shapes are concerned, even if participants with low vision perceived no interference for the understanding of the shapes, blind participants required additional information and all participants coincided in the fact that such faces did not carry meaningful information. The participants also showed some confusion with the size of the shapes in the frames. Since in the second version of the T-SAM the size increases progressively, similar shapes could be easier to relate. Participants with low vision suggested including augmented numbers apart from the numbers in braille.

The final task concerning the representation of the dimension of arousal invited participants to provide other ways to represent the grading. A solution was offered based on the use of different materials to convey different textures, from low arousal, related to the state of being calmed, to a high arousal, related to the state of being excited. A scale could be portrayed from a silkier texture, using materials such as cotton or linen, to a coarser texture, such as that of the rope or the sandpaper. The application of such a strategy for the representation of scales needs to be further investigated in research. It was also suggested by participants that touch is also subjective and that not every person would relate a texture with the same meaning.

To sum up, the main recommendation acknowledged in this phase was the removal of the facial features inside the shapes and the inclusion of more textured surface in the final image of this series, as well as the representation of all the shapes with the same size rather that in a progression. Such recommendations were implemented in the last version of the T-SAM (Figure 21). An explanatory video on the creation of the T-SAM con be visited online in the TransMedia Catalonia research group YouTube page (TransMedia, 2019).

The T-SAM and its Two Versions

In the last phase participants were invited to discuss their whole understanding of the questionnaire when comparing the two versions. In general, all participants agreed that once the explanation was given to them and they could interact with the questionnaire, they understood their task and the structure of the two dimensions.

All preferred the second version. The explanation can be found in the physical characteristics of the relief between one and the other. The first version of the questionnaire T-SAM (Figure 19) was made with fewer layers of gloss for the relief, while the second (Figure 20) had a better-defined relief. This was appreciated by all the participants whether low-vision or blind. Apart from the more technical characteristics, they preferred the outline of the faces and the removal of the faces inside the shapes in the second dimension.

The changes suggested by the participants in the focus groups were then implemented in the second version of the T-SAM (Figure 20). The final result is shown in Figure 21.

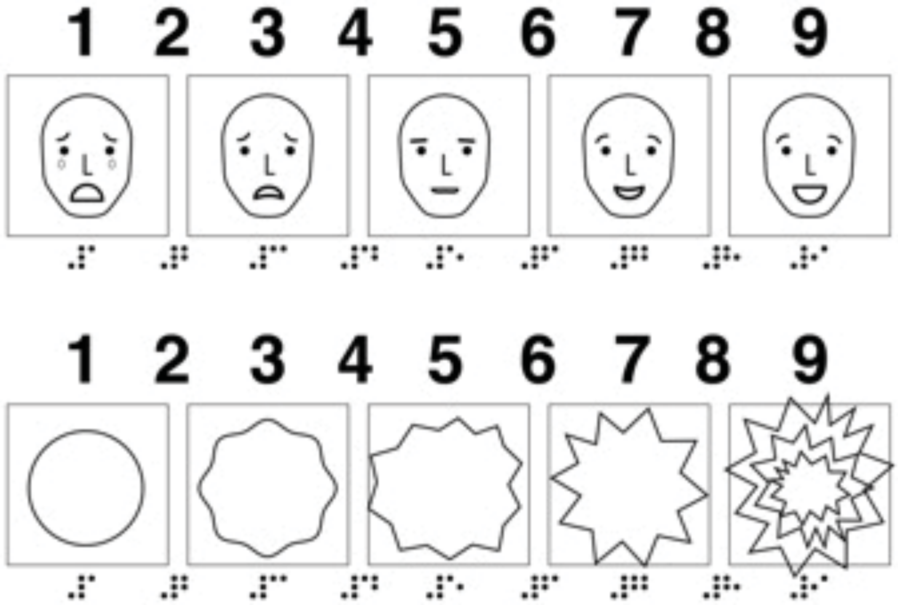


Figure 21. Final version of the T-SAM

Finally, all participants acknowledged that even if some of the characteristics were less useful for people with certain specific needs (emoticon-like features of the faces for blind participants), the questionnaire could be used by people with different visual conditions, with the incorporation of the augmented numbers for people with low vision as well.

2.3. Experiment 2

The previous paragraph described the development of the experiments performed with audio subtitled stimuli, which conformed the so-called Experiment 1. The following section describes Experiment 2, carried out with sighted participants and subtitled clips as stimuli.

2.3.1. Duration

Experiment 2 had a total duration of around 45 minutes. As in the case of Experiment 1, the main part of the experiment had a duration of 30 minutes. In this main body of the experiment the stimuli were presented by the procedure control software, which also provided the instructions to the participants. To this 30 minutes the time of explanation, the time for the connection of the sensors, and completion of self-report instruments were added. The time in Experiment 2 is reduced in comparison to Experiment 1, due to two factors. First, questionnaires are completed by the participant directly on the computer, not orally as in Experiment 1. And second, this experiment does not finish with a questionnaire on preference.

2.3.2. Participants

Experiment 2 was performed with 42 participants: 27 female and 15 male aged between 21 and 63 years old (median=24). The group of sighted participants were contacted through social media and other kind of advertising, such as mailing lists and networking tools.

2.3.3. Materials

In this section the materials used in Experiment 2 are described. First of all, the technological apparatus is presented, including the recording and displaying devices. Afterwards, the self-report instruments are introduced according to the moment throughout the experiment in which they were administered.

2.3.3.1. Apparatus

Psychophysiological data were collected by means of the Captiv L7000 (Figures 11, 12 and 13) and were controlled by the procedure software of Tobii X20-60 (Figure 14) as it was done in Experiment 1 (see Section 2.1.). Both devices were connected to the computer from which the presentation of stimuli was monitored just like in Experiment 1.

A second computer, just like in Experiment 1, was used for the registration of the questionnaires' answers. The same laptop MacBook Air was used. The computer was connected to the internet and Google Forms sheets were used for every questionnaire. In this case, however, the participant was interacting with the computer directly.

2.3.3.2. Instruments

Some of the self-report instruments used in Experiment 2 were the same ones used in Experiment 1, however there were some changes. The reason being that Experiment 2, performed with sighted participants, was also part of another project that used the same methodology and stimuli but aimed to collect other kind of data. Taking into consideration this fact, the self-report instruments administered to the participants were the following.

Consent Form

This consent form has been approved by the ethical committee of the Universitat Autònoma de Barcelona (see Section 2 in the Annexes). The consent form stated that any personal details or any other data from the participants will be confidential and that the participant could stop the experiment at any moment without consequences.

Pre-Questionnaire

The pre-questionnaire included:

a) Demographic questions
b) Questions about anxiety and caffeine, drug and medicament consumption (see Section 2 in the Annexes).
c) the Toronto Alexithymia Scale (TAS-20) (Taylor et al, 1985) to test their ability to identify and describe emotions in the self. Different translations of the TAS-20 questionnaire into Spanish are found. For the present study it was decided to use the validated Spanish translation in Martínez Sánchez (1996). The contents of the Alexithymia Scale can be found in the Annex of this thesis (see Section 2 in the Annexes).
d) As participants were also part of another experiment (in parallel with this project), they had to respond to the State-Trait Anxiety Inventory (STAI) (Spielberg, Gorusch, Lushene, Vagg, & Jacobs, 1983). It has two different versions, the state and the trait forms. The state form aims to measure the anxiety of a person in the moment the questionnaire is completed and it is compared to the outcome in the trait form, which centres on the general tendency of the person towards anxiety. Only the second form, the trait form, was administered in this experiment.

T-SAM Questionnaire

The questionnaire administered in Experiment 2 was the same T-SAM explained in Experiment 1. The use of the same version of the SAM questionnaire provides better cohesion when contrasting results between sighed, and blind and partially sighted participants.

Final Questionnaire

As a part of this parallel experiment, a final questionnaire was administered, the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ) (Marvin Zuckerman, 2002) in its shortened version (ZKPQ-50-C) (Aluja et al., 2006). It is a general questionnaire on personality divided in 5 subscales: Impulsive Sensation Seeking (ImpSS), Neuroticism-Anxiety (N-Anx), Aggression-Hostility (Agg-Host), Activity (Act), Sociability (Sy). The most relevant scale for the purposes of this experiment is the ImpSS, which is related to emotional perception.

2.2.3.2. Stimuli

The clips used in Experiment 2 corresponded to the same scenes as in Experiment 1 but in a subtitled version. The translation was done by a professional translator into Spanish (see Section 2.1.3.3).

2.2.4. Procedure

All participants are exposed to the 3 scenes in a single condition: subtitled in Spanish. The order of presentation of scenes was randomised by the experimental procedure software (Tobii). The same protocol was repeated for the 42 participants.

2.2.4.2. Experiment Outline

1. Participant enters the laboratory.
2. Participant is given the information sheet and is informed about the procedure and where the sensors are going to be placed.
3. Participant is given the consent form. The participant signs the consent form and is informed that they can withdraw at any point of the experiment.

4. The sensors are connected to the participant and the software is started.
5. The participant completes the demographic and general questions on a Google Forms sheet.
6. The participant completes the TAS-20 questionnaire on a Google Forms sheet.
7. The participant completes the STAI questionnaire on a Google Forms sheet.
8. The participant is then provided with headphones. From that moment, the presentation guides the participant by means of instructions appearing on screen and recorded orally by the researcher.
9. The participant is first invited to a relaxation period induced by mood music, which lasts 8 minutes. This relaxation period is aimed at unifying the participants' emotional state and to record the base line.
10. The first clip is played.
11. The participant completes the T-SAM.
12. The participant is invited to a relaxation period (3 minutes) before the next clip. This is repeated two more times.
13. The participant competes the ZKPQ-50-C questionnaire on a Google Forms sheet.
14. The participant can remove the headphones and disconnect the sensors.
15. The participant is thanked for their collaboration.

2.2.5. Pilot Experiment

The protocol was tested with four sighted participants. The duration of the experiment was expanded since it was slightly longer than expected. No important issues arose, and the protocol was maintained without changes. The results obtained in the pilot participants were also used for analysis, since no changes were made to the protocol and they were valid.

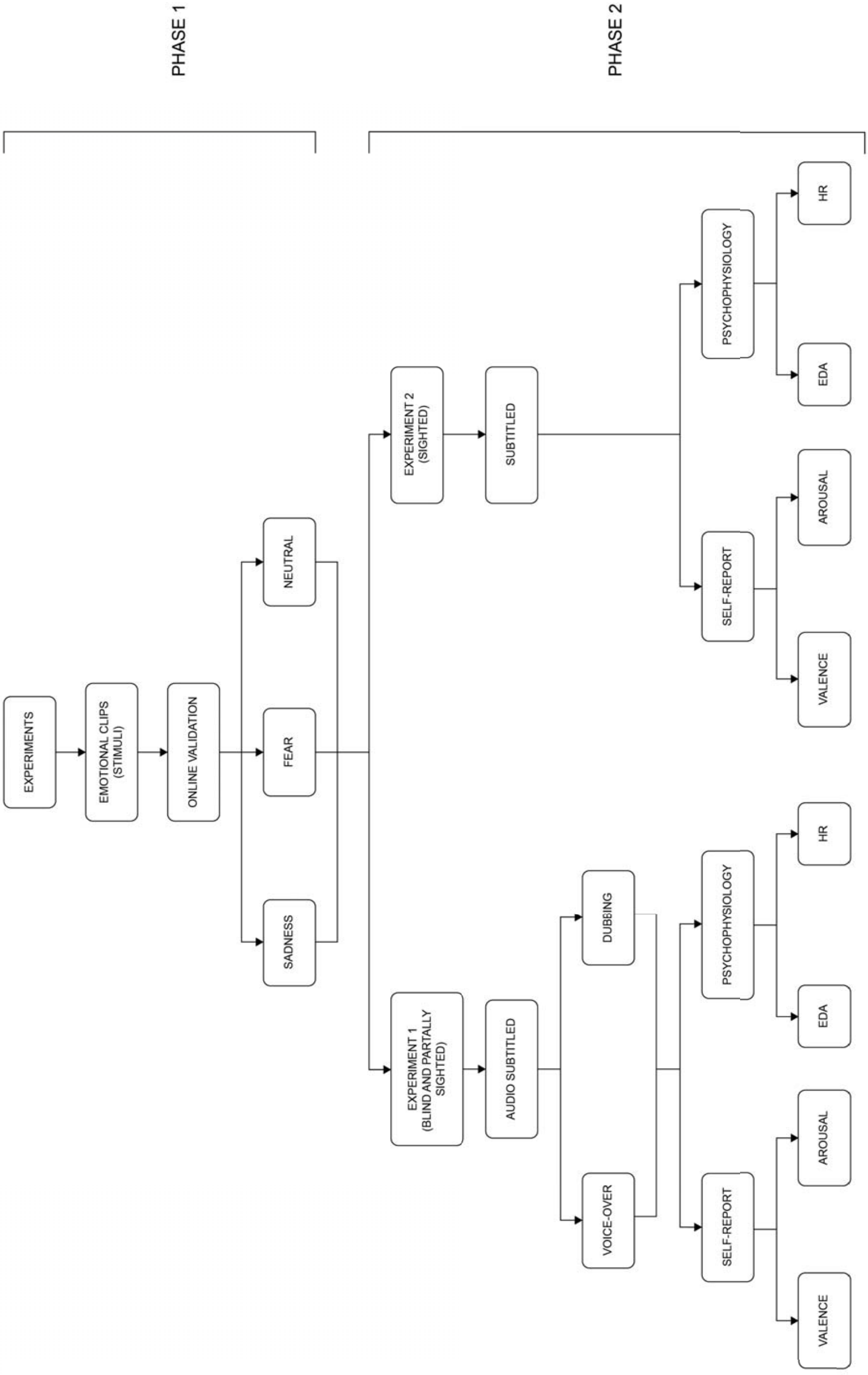


Figure 22. Structure of Study 2

3. Data Analysis

The same measuring methods were used in both Experiments 1 and 2. The methodology followed to analyse the data will also be the same. This section will be divided into two main parts: the analysis of data for the self-report questionnaires (T-SAM) and psychophysiological measures (EDA and HR).

3.1. Self-Report Measures: T-SAM

In both Experiments 1 and 2, the T-SAM questionnaire was administered to the participants. Each participant provided a number from 1 to 9 for valence and arousal respectively. The results are presented in both median and averages (with standard deviation). The analysis was conducted through tests for non-parametric data. The tests used were the Related-Sample Friedman's Two-Way Analysis of Variance by Ranks when only one variable was considered (emotion or AST effect) and a General Linear Mixed Model for Repeated Measures when two variables were considered (emotion and AST effect).

For the comparison of the results in both experiments pairwise comparisons were conducted with the Whitney-Mann test across the three emotions for both AST effects and between both groups (blind and partially sighted, and sighted)

3.2. Psychophysiological Measures

The data were collected by the software of the device, the Captiv L7000 at a frequency of 32 Hz. A file saved under the format .cvs was extracted for the recording of each participant. The recordings had a duration of around 30 minutes each and had to be pre-processed in order to analyse them. Each 30-minute recording was complemented with synchro points that indicated the beginning of the recording, the beginning and the end of each stimuli and the end of the recording. The .cvs files were transformed into .xlsx and processed with Excel at this step. Each file presented three columns: one with the time code and two others with the value (for EDA and HR respectively). Excel files with the raw data for the participants are found in the Electronic Annexes.

Each .xlsx file was treated with a labelling proces that aimed to mark the duration of the stimuli, with their emotion and effect, and the relaxation period from which the baseline was going to be extracted. The labelling process was conducted with Excel. A specific code was created for the labels in the data in a way that the analysis could easily identify the values to be processed, e.g. 3 for sadness, 4 for neutral, and 5 for fear; 1 for dubbing, 2 for voice-over, etc. Once each of the parts of the recording was identified, the values of the stimuli were analysed in terms of averages minus the average of the relaxation period (baseline). As baseline, it was decided to use the last 3-minute period of the initial induced-relaxation time. Once the Excel files were correctly labelled, data were analysed with a statistical analysis software, the Statistical Package for the Social Sciences (SPSS).

3.2.1. Electrodermal Activity

Data were collected by means of a sensor placed in contact with the skin on the second phalange of the pointing and middle fingers of the non-dominant hand. The values are expressed in micro Siemens (μS) and show the difference between the average per stimuli and the baseline recorded during the last three minutes of the relaxation period. The results were statistically analysed with General Linear Model for Repeated Measures analyses. For the comparison of both variables (emotion and AST effect) a General Linear Mixed Model for Repeated Measures analysis was conducted. When comparing both groups (blind and partially sighted, and sighted), T-tests were conducted to test the differences.

Due to technical problems during the recording of EDA, 9 participants were discarded for this measurement in Experiment 1 and 6 participants were discarded in Experiment 2. This fact caused the total number of participants to be 33 in Experiment 1 and 36 in Experiment 2.

3.2.2. Heart Rate

The measurement of HR carried out in this experiment was based on the changes in the beats per minute (bpm) for each participant. In order to analyse the data, the average of the bpm for each stimulus was calculate in the relation to the baseline consisting of the three last minutes of the relaxation period at the beginning. Data were collected by means of a sensor in the form of a belt that was placed in contact with the skin right below the chest.

A total of 9 participants from Experiment 1 and 11 participants from Experiment 2 were excluded from the analysis. This left the analysis of HR with 33 participants in Experiment 1 and 31 participants in Experiment 2.

Chapter 7

Discussion of Results in Study 2

Chapter 7
Discussion of Results in Study 2

This chapter discusses the results of Study 2, which includes Experiments 1 and 2. The two experiments are performed by following the same protocol and using the same content treated in a different way depending on the participants they were exposed to.

Experiment 1 presents audio described and audio subtitled clips (with a voice-over and a dubbing effect) to blind and partially sighted participants. Results related to Experiment 1 are presented and discussed for the self-report instrument, T-SAM (Tactile-Self Assessment Manikin), (**Section 1.1.**); for the psychophysiological measures, namely EDA and HR (**Section 1.2.**); and for the questionnaire on preferences (**Section 1.3.**).

Experiment 2 presents the same contents in their subtitled version to sighted participants. For Experiment 2 results are shown and discussed for the self-report instrument, T-SAM questionnaire (**Section 2.1.**), and for the psychophysiological measures, EDA and HR (**Section 2.2.**). A preference questionnaire is not administered in Experiment 2 since there are no AST effects that could be compared.

Finally, **Section 3** deals with the comparison of the results between both Experiment 1 and Experiment 2. The results are compared following the same order in which they are presented: first, the two dimensions of the T-SAM questionnaire, then the results for EDA and HR in the part of psychophysiological measures. In **Section 4**, the correlation of the results obtained through psychophysiological measures and those obtained by means of the self-report instrument are discussed.

Common to all the participants in both Experiment 1 and Experiment 2 is the pre-questionnaire, which consists of some demographic questions and other questions related to the levels of anxiety, the consumption of medicines and the consumption of drugs; and a questionnaire on alexithymia, called TAS-20, which is intended to test the ability of the participants to regularly understand and express emotions. These pre-experimental questions are only aimed to prove the suitability of the participants for the experiments.

1. Experiment 1

In this section, the results of the Experiment 1 are discussed. The results are presented in a general order: first the data collected through self-report instruments, second the results obtained through psychophysiological measures and finally, the results obtained from the questions on preference.

Before beginning, it is worth stating that results in the TAS-20 questionnaire did not did not prompt the removal of any participant since the results were considered as normal for the 42 participants. A person is considered to show signs of alexithymia from the threshold of 50 points (out of 100), and the degree increases with the number of points.

1.1. Results for the Self-Report Questionnaire

The T-SAM is divided into two dimensions: valence and arousal (see Chapter 3, where emotions and their dimensions are discussed). Both dimensions were rated by means of a 9-point scale. Valence provides information about the emotion being positive or negative. In the scale, 1 is very negative and 9 very positive. For arousal, which provides information about the activation (arousal) caused by the emotion, 1 meant that the person was unaltered by the scene and 9, that the person was very excited or activated. The results are shown in average marks for each of the dimensions.

1.1.1. Valence

Figure 23 shows the median and average (with the standard deviation) of the ratings for the valence dimension in Experiment 1 without taking into consideration the effect (voice-over and dubbing). The vertical axis shows the rating from 1 to 9, the horizontal axis shows the three emotional clips (for the emotion of sadness, fear and neutral, respectively).

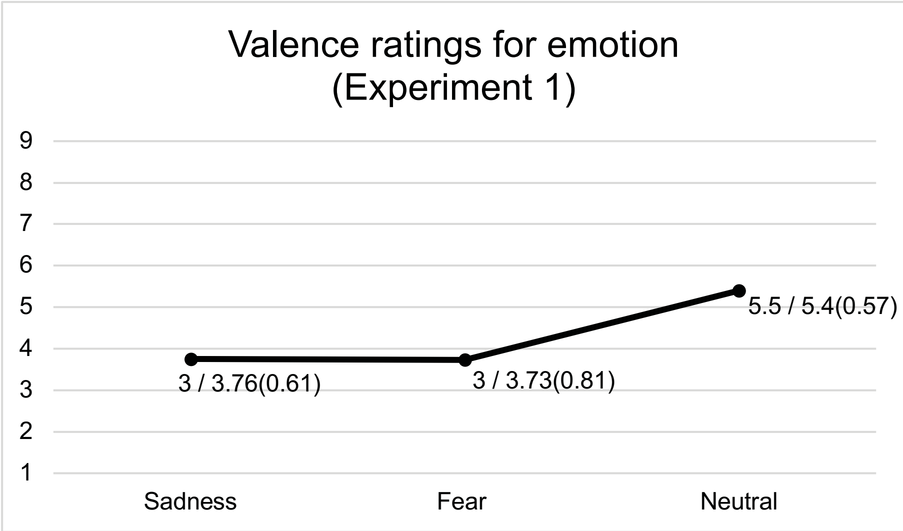


Figure 23. Valence ratings for emotion (Experiment 1): median, average (standard deviation)

To check for the differences between the ratings of the emotions, non-parametric tests were used since the ratings are not a continuous variable. The Related-Sample Friedman's Two-Way Analysis of Variance by Ranks test was carried out. The test yielded statistically significant differences among the emotions ($Q_{(2)} = 24.700$, $p < .001$). To further investigate the differences among emotions, a pairwise comparison was conducted. The comparison showed significant differences between the rating of sadness and neutral clips ($Q_{(1)} = 25.000$, $p < .001$) and between neutral emotion and fear-inducing clips ($Q_{(1)} = 9.529$, $p < .001$). There was no significant difference between the ratings of sadness- and fear-inducing clips ($Q_{(1)} = 1.581$, $p = .209$).

It can be argued that the emotions targeted were satisfactorily induced in terms of valence according to the literature (see Chapter 3). The clips which were intended to induce the emotions of sadness (3.76) and fear (3.73) obtained lower rates in comparison to the clip portraying a neutral emotion (5.4). The sadness and fear clips have been rated with less valence, that is, more negative, in comparison to the neutral clip, rated as more positive. The statistical analysis performed with a Related-Samples Friedman's Two-Way Analysis of Variance by Ranks. It pointed out that the statistically significant differences were found between the ratings for the sadness clip and the neutral clip, and the fear clip and the neutral clip. No significant differences are found between the ratings of sadness and fear clips, which was an expected result since both emotions are both typically low in valence.

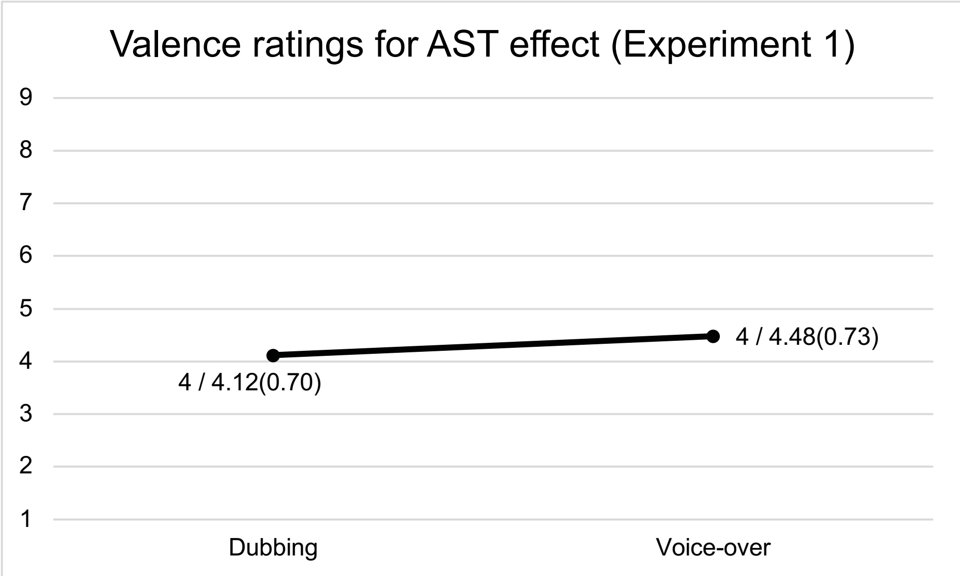


Figure 24. Valence ratings for the AST effect (Experiment 1): median, average (standard deviation)

Regarding the AST effect, the Related-Samples Friedman's Two-Way Analysis of Variance by Ranks showed that the participants' valence ratings did not significantly differ depending on the way the audio subtitle was presented ($Q_{(1)}$ = 1.778, p = 1.778) (Figure 24).

The results for the ratings of valence for the condition of effect were also obtained by means of calculating the average of the ratings of the 42 participants of the experiment. Only by observing the effects and ignoring the variable of emotions, it is shown that the clips presenting audio subtitles with a voice-over effect were rated slightly higher in valence (4.48) in comparison to clips with a dubbing effect (4.12).

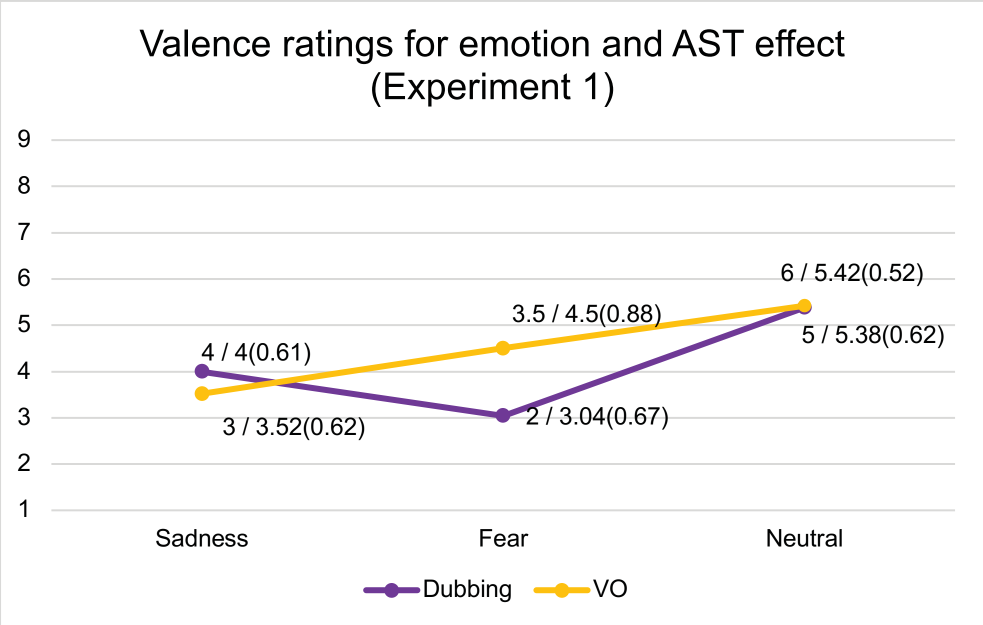


Figure 25. Valence ratings for emotion and AST effect (Experiment 1): media, average (standard deviation)

Figure 25 shows the median and average results for the dimension of valence for both AST effect. Since non-parametric tests do not allow for the analysis of two variables at the same time, in this case a General Linear Mixed Model for Repeated Measures analysis was applied in order to test the differences between both effects across emotions.

The AST effect (dubbing and voice-over) only proved to be different for the clips inducing the emotion of fear (p = .01), confirming that dubbing caused a more negative effect. The statistical analysis does not yield significant differences between AST effects for clips portraying the emotion of sadness (p = .59) and a neutral emotion (p = .72).

Overall, clips with a voice-over effect were rated higher in valence in the clips portraying fear (3.52) and a neutral emotion (5.42) in comparison to the dubbing effect which shows lower valence in fear (3.04) and neutral (5.38). Contrarily, for the sadness-inducing clips, the ratings are higher for the audio subtitles with a dubbing effect (4) in comparison to those with a voice-over effect (3.52). Even if dubbing and voice-over effects show almost inexistent difference for the clips portraying sadness and a neutral emotion, there is a difference of AST effects in the ratings of the clips conveying fear. Fear is defined as an emotion with low valence, i.e. a negative emotion. Therefore, a dubbing effect in the audio subtitles could indicate that the participants experienced a more negative emotion.

1.1.2. Arousal

Figure 26 shows the median and average results for the second dimension of the T-SAM questionnaire: arousal, and presents the interaction between the emotions: sadness, fear, and neutral.

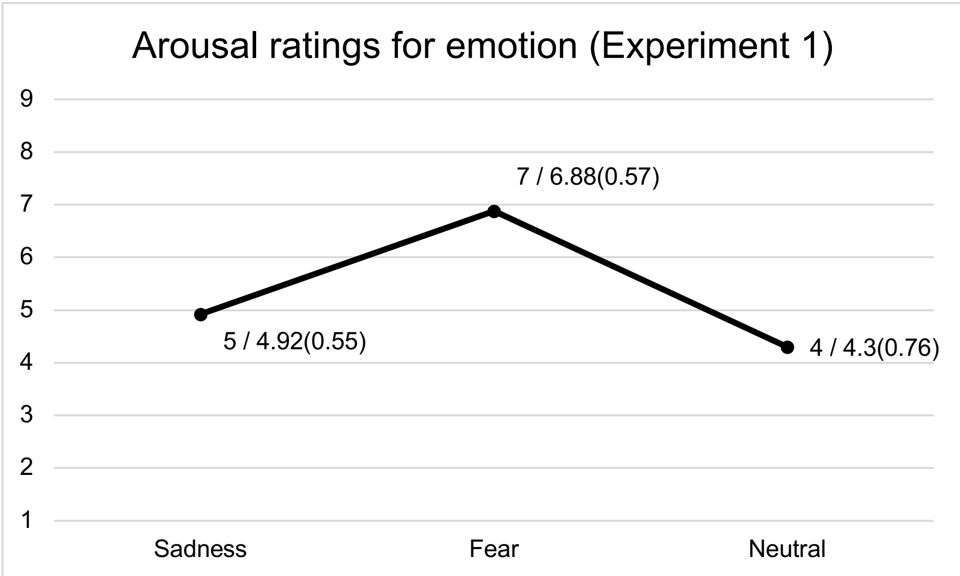


Figure 26. Arousal ratings for the emotion (Experiment 1): median, average (standard deviation)

In order to compare the ratings for the condition of emotion the Related-Samples Friedman's Two-Way Analysis of Variance by Ranks was carried out. The test shows that there are statistically significant differences among the emotions ($Q_{(2)}$ = 41.367, p < .001). A pairwise comparison was performed to check the differences among the three emotions. The comparison between sadness and fear proved to be significant ($Q_{(1)}$ = 23.684, p < .001) as well as that between neutral and fear ($Q_{(1)}$ = 28.900, p < .001). There was no significant difference between sadness and neutral ($Q_{(1)}$ = 3.789, p = .05).

The emotions were induced in terms of arousal accordingly to the literature (see Chapter 3). The clips aiming to induce fear were rated with higher arousal (6.88) in relation to clips inducing sadness (4.92) and a neutral emotion (4.3). A Related-Samples Friedman's Two-Way Analysis of Variance by Ranks confirmed the initial observation and the differences were statistically significant ($p < .001$).

A pairwise comparison revealed that the differences were significant only for the comparison between sadness and fear ($p < .001$) and fear and neutral ($p < .001$). It could be stated that fear caused more activation in the participants in comparison to sadness, considered to obtain low-medium arousal and neutral, which was expected to cause the lowest of the arousal ratings.

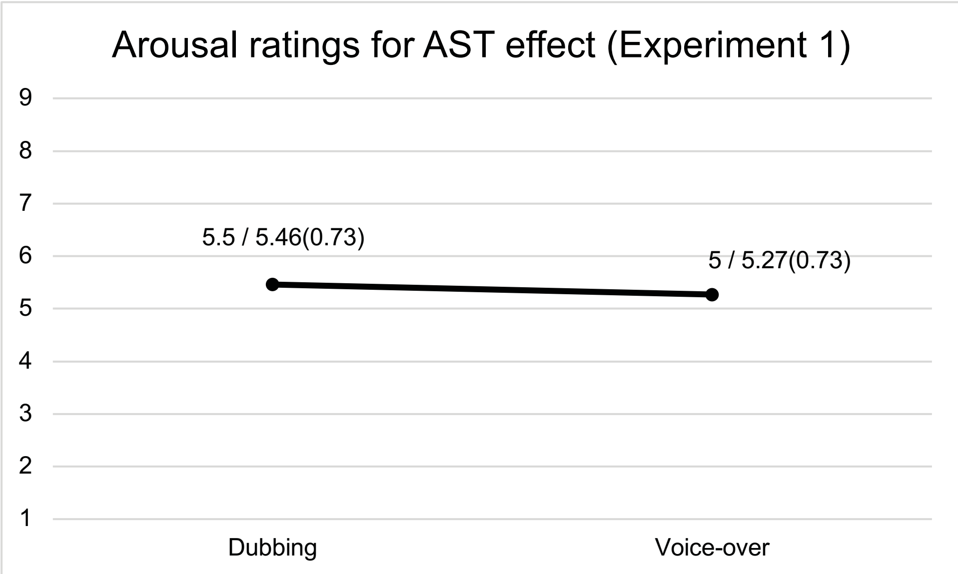


Figure 27. Arousal ratings for the AST effect (Experiment 1): median, average (standard deviation)

As far as the effects were concerned, the Related-Samples Friedman's Two-Way Analysis of Variance by Ranks shows that there is no significant differences between the two AST effects ($Q_{(1)} = 1.778$, $p = .18$) (Figure 27). The averages of the ratings for arousal taking into consideration the type of effect seem very similar, with the dubbing effect clips obtaining a slightly higher rating (5.46) in comparison to the voice-over effect (5.27).

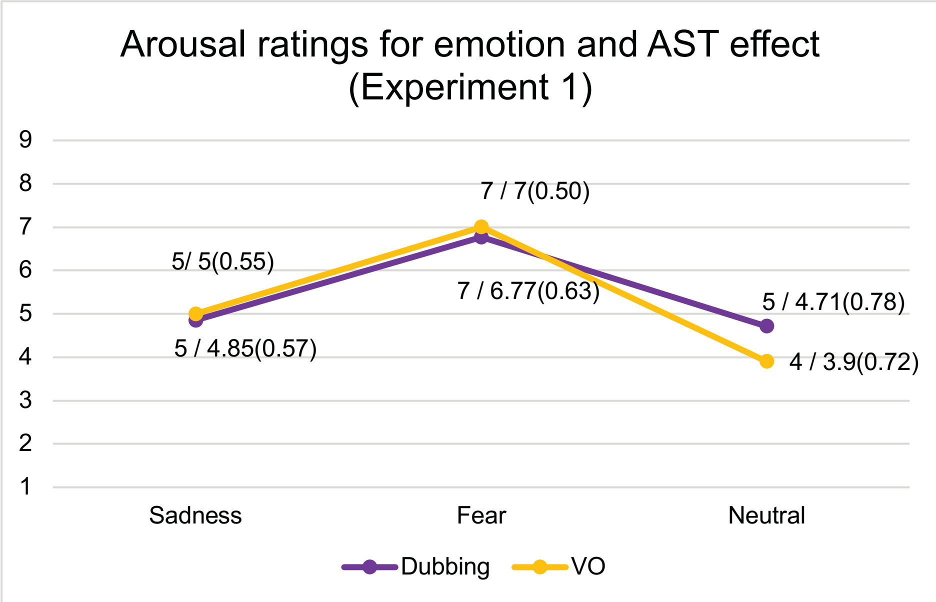


Figure 28. Arousal ratings for emotion and AST effect (Experiment 1): median, average (standard deviation)

Figure 28 shows the median and average ratings in arousal for AST effects across emotions. In order to test the differences, a General Linear Mixed Model for Repeated Measures was conducted. The AST effect does not show any significant differences for any of the emotion-inducing clips: sadness-inducing clip (p = .80), fear-inducing clip (p = .63) and the clip aiming for a neutral emotion (p = .33).

The comparison of the two effects across the three emotions did not show any significant difference. As for the trends observed, it is only worth remarking the difference observed in the averages of the ratings for the neutral clips. The dubbing effect has been rated as higher in arousal (4.71) in comparison to the voice-over effect (3.9). This could be read as sharing a similar idea of the results obtained in valence, that a dubbing effect causes more activation.

1.2. Results for Psychophysiological Measures

In this section, the results obtained for the measurements of EDA and HR are presented. The results, expressed in graphs, are accompanied by the statistical analysis.

1.2.1. Results for EDA

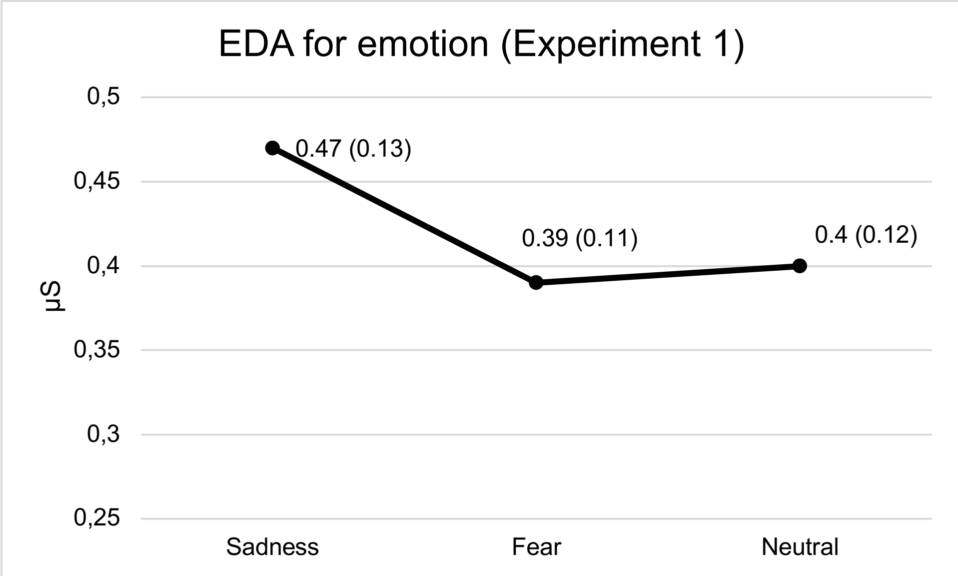


Figure 29. EDA stimuli averages minus 3-minute baseline average for emotion (Experiment 1)

Figure 29 shows the averages of the values for EDA minus the baseline. In order to perform EDA data analyses, the average for the values obtained with EDA recording during each of the clips (emotions) was calculated. Later the difference between the average values for each clip minus the average of the baseline (which consisted of the last 3 minutes of the EDA recording during the initial period of relaxation) was calculated.

Since this measure is continuous, a General Linear Model of Repeated Measures analysis was applied to these values. The differences do not reach statistical significance (F (2, 64) = .21; p = .81, Partial Eta Squared = .007). The values (see Figure 29), which consisted of the stimuli averages minus the 3-minute baseline average, were very similar: sadness (0.47 μS), fear (0.39 μS) and neutral (0.4 μS). Therefore, there are no remarkable trends.

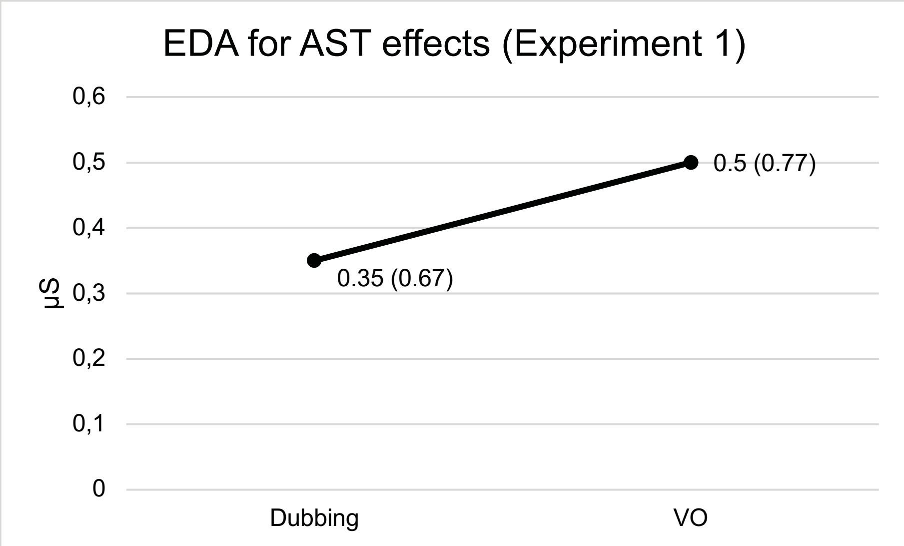


Figure 30. EDA stimuli averages minus 3-minute baseline average for AST effect (Experiment 1)

Although the voice-over obtained higher values (0.5) than that of dubbing (0.35) (Figure 30), the General Linear Model of Repeated Measures shows that the differences between the two AST effects are not statistically significant (F (1, 32) = 0.32; p = .57; Partial Eta Squared = .01). However, a small difference can be observed between the measures obtained during the voice-over effect (0.5 µS) and dubbing effect (0.35 µS). This would imply that the voice-over effect has caused higher EDA activation.

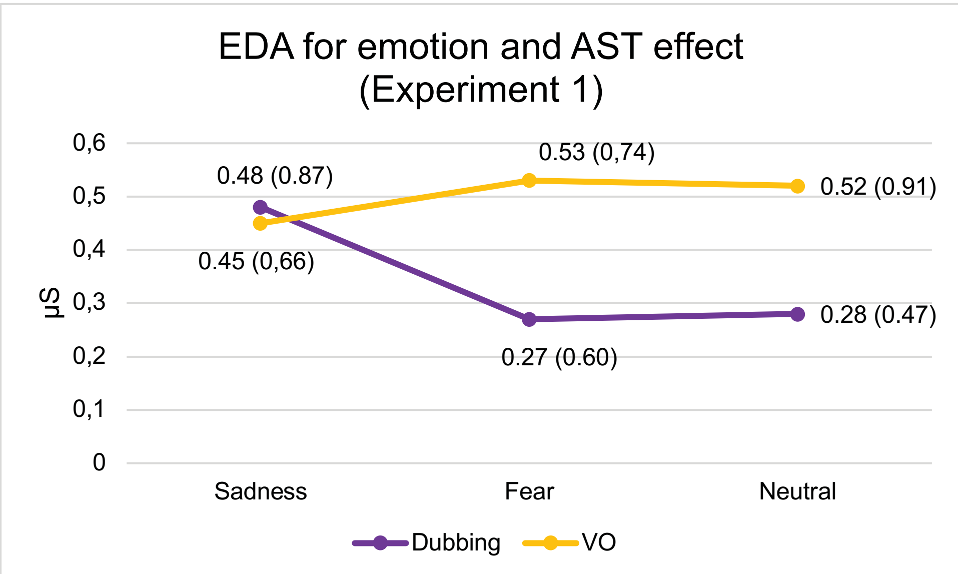


Figure 31. EDA stimuli averages minus 3-minute baseline average for emotion and AST effect (Experiment 1)

Figure 31 shows the interrelation between the two conditions, emotion and AST effect. The General Linear Mixed Model of Repeated Measures did not report any significant differences in the pairwise contrasts between the two AST effects through the three emotions: sadness (p = .45), fear (p = .52) and neutral (p = .17).

Across emotions some trends can be seen. Apparently, the effect of voice-over effect produced higher values for fear (0.53 μS) and neutral (0.52 μS) emotions in contrast with a dubbing effect (0.27 μS and 0.28 μS, correspondingly). In sadness, the average values are very close for the two AST effects (dubbing = 0.48 μS and voice-over = 0.45 μS). However, the results are inconclusive, and more experiments are required on the topic.

1.2.2. Results for HR

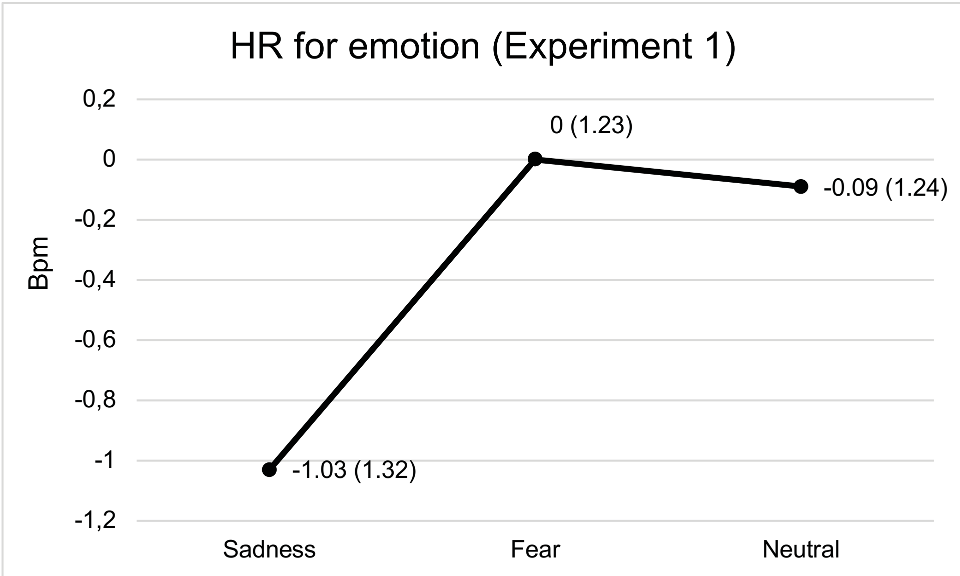


Figure 32. HR stimuli averages minus 3-minute baseline average for emotion (Experiment 1)

The General Linear Model of Repeated Measures analysis provided significant differences for HR in terms of emotion (F (2, 64) = 3.320, p = .04; Partial Eta Squared = .94).

The values obtained with the sadness-eliciting clip are the lowest (-1.03 bpm) and are different from neutral (-0.09 bpm) and very close to be significantly different from those resulting from the fear-inducing clip (0 bpm) (Figure 32). Pairwise analyses indicate that the values obtained during the recordings for the clips portraying sadness and neutral emotions are different ($t_{(32)}$ = -2.40; *p* = .021); and differences between those portraying sadness and fear almost reach significance ($t_{(32)}$ = -1.973, *p* = .057).

According to these results, HR presents more deceleration during the sadness-inducing clip, which correlates to the information provided in the description of emotions in Chapter 3. The results, though, are all negative, meaning that HR is slower during the stimuli. This could be related to concentration, that is proven to cause deceleration (Graham, 1992; Petrie Thomas, Whitfield, Oberlander, Synnes, & Grunau, 2012; Tremayne & Barry, 2001).

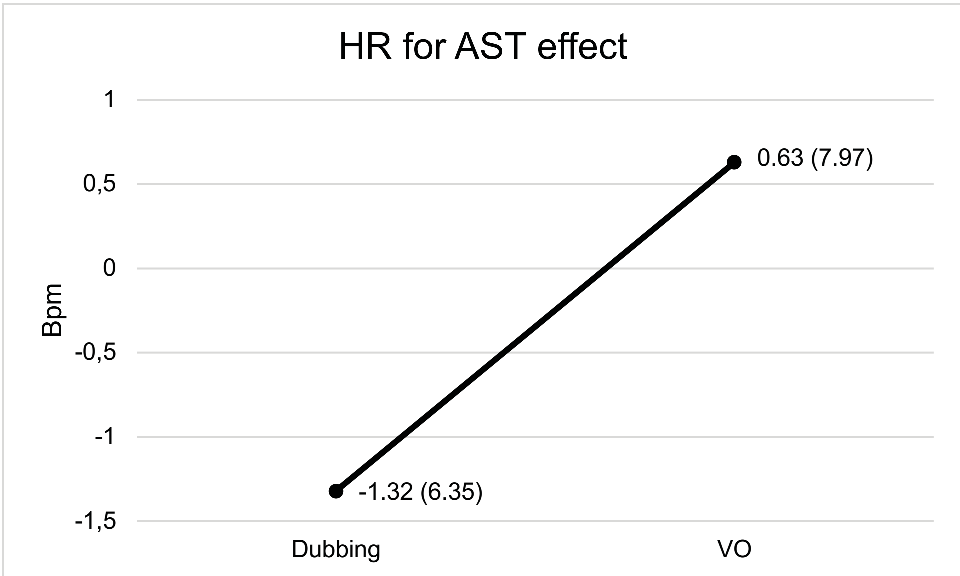


Figure 33. HR stimuli averages minus 3-minute baseline average for AST effect

In Figure 33, a difference might be observed between the voice-over effect (0.63) and the dubbing effect (-1.32). However, after applying a General Linear Model of Repeated Measures, the results did not yield significant differences (*F* (1, 32) = 0.88; *p* = .35; Partial Eta Squared = 0.027).

By observing the differences obtained in the calculation of the values (stimuli averages minus 3-minute baseline average) for the two effects, a voice-over effect results in more acceleration (0.63 bpm) in comparison to dubbing effect (-1.32 bpm) which result in deceleration. This could be also linked to the process of concentration mentioned previously, suggesting that the dubbing effect triggers more concentration in the participants, but results are not conclusive.

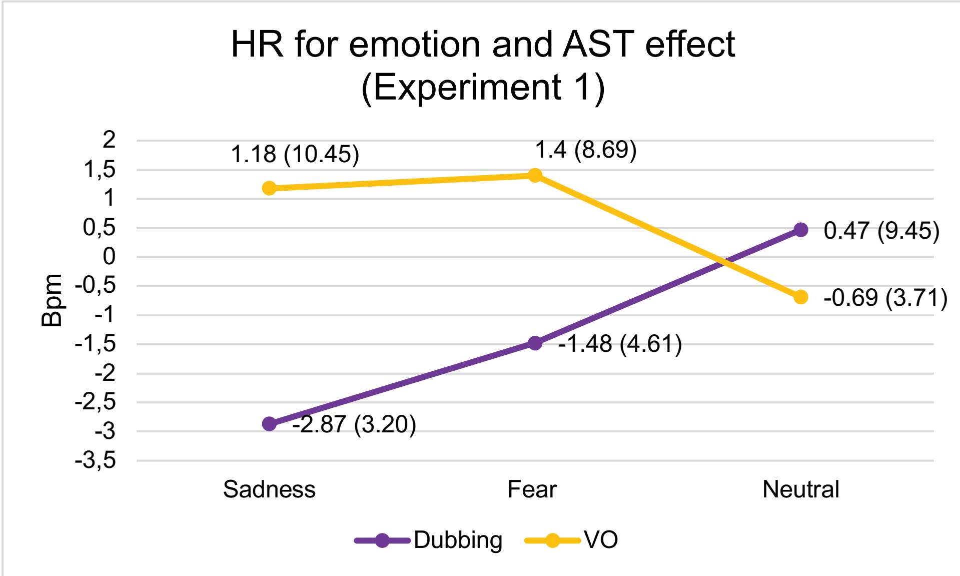


Figure 34. HR stimuli averages minus 3-minute baseline averages for emotion and AST effect (Experiment 1)

Figure 34 shows the interrelation between the two conditions, emotion and AST effect for HR. The General Linear Mixed Model of Repeated Measures did not report any significant differences for any of the pairwise comparisons of the AST effects through emotions: sadness (p = .12), fear (p = .24) and neutral (p = .64).

1.3. Results for the Preference Questionnaire

This questionnaire aimed to collect some qualitative and subjective information from the participants of the experiment (see Annex 2). It was divided in two parts: audio subtitling and audio description. All the questions in the questionnaires contained a “Why” question so participants could provide explanations to their answers if they wished to do so.

Before presenting this questionnaire to the participants it was explained to them that two different strategies were used to voice the subtitles and that the AD had the same characteristics in the three videos. This allowed for the participant to think about what they had heard and to compare the stimuli they had been exposed to.

1.3.1. Audio Subtitling

The first part of the questionnaire included 6 questions related to the different characteristics of the audio subtitles and how such characteristics had an impact on the preferences of the participants, their comprehension of the scenes and on how emotions were conveyed.

Which one of the two options do you prefer?

[¿Cuál de las dos opciones prefieres?]

This first question aimed to distinguish the preferences of the participants in terms of the effect of the ASTs, i.e. dubbing and voice-over effect. The question includes three different possible answers:

- The less acted voice in which the original can be heard [La voz menos interpretada en la que sí se oye el original]
- The more acted voice in which the original cannot be heard [La voz más interpretada en la que se oye el original]
- It does not matter [Me es indiferente]

In Figure 35 the results for this first question are shown. Only two out of 42 (5%) responded that they did not have a preference for any of the two effects. The answers for the two options related to the two effects are quite balanced: 21 participants (50%) preferred the more acted voice without the original being heard, thus related to the dubbing effect, while the remaining 19 (45%) preferred the less acted voice superimposed to the original dialogue, related to the voice-over effect. The reasons behind these preferences varied depending on the interest of the participants, some of them were accustomed to audiovisual content in foreign languages and enjoyed the original language and acting.

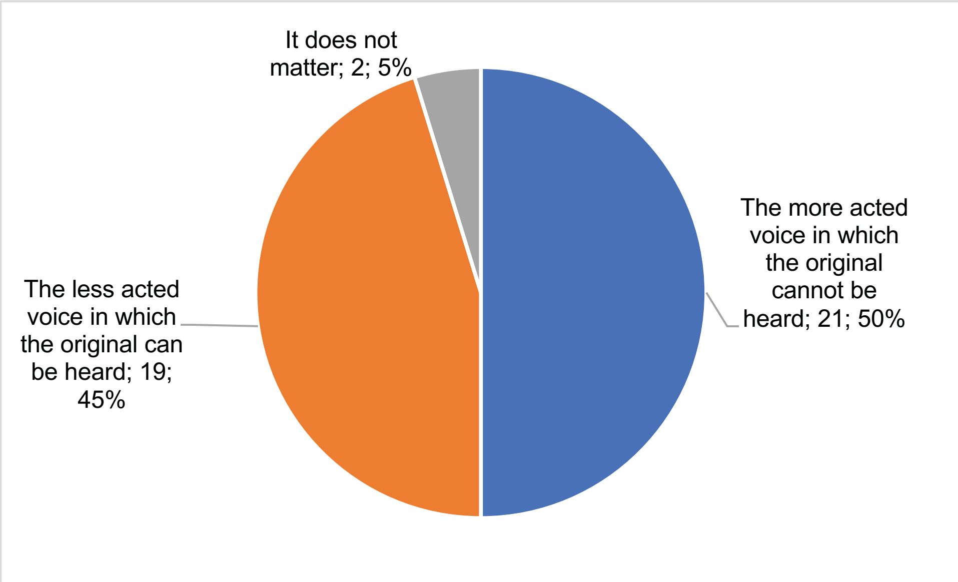


Figure 35. Replies to “Which one of the two options do you prefer?”

Which one do you think can transmit better the emotion of the scene?

[¿Cuál consideras que permite transmitir la emoción de la escena?]

This second question was intimately related to the previous one, which is reflected in the answers. As shown in Figure 36, again, three answers were possible:

- The less acted voice in which the original can be heard [La voz menos interpretada en la que sí se oye el original]
- The more acted voice in which the original cannot be heard [La voz más interpretada en la que se oye el original]
- Both transmit emotions in the same way [Ambas transmiten la emoción del mismo modo]

Only two people (5%) answered that in both cases the emotions were transmitted in the same way. Of the remaining participants, 18 (43%) suggested that a less acted voice where the original dialogue can be heard, thus more related to the voice-over effect, transmits emotion in the most effective way, while the majority, 22 (52%) responded saying that a more acted voice in which the original is not heard (i.e. dubbing effect) is the best way to transfer emotion.

Even if the difference is not very important, the preference within the group of participants was the dubbing effect.

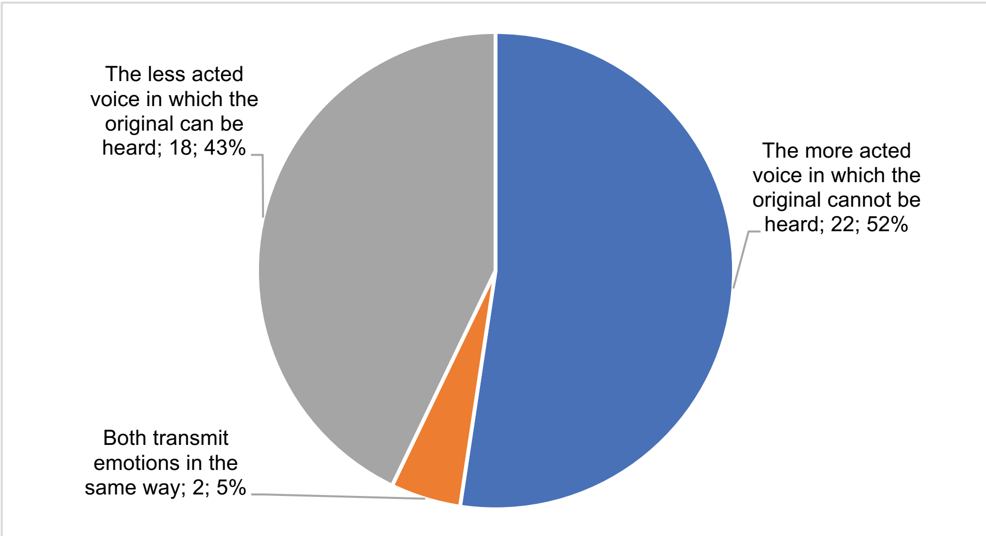


Figure 36. Replies to “Which one do you think can transmit better the emotion of the scene?”

Was there an option in which dialogues where more difficult to understand?

[¿Ha habido una de las dos opciones en las que ha sido más difícil comprender los diálogos? ¿Cuál?]

This question aimed to collect answers on the self-reported comprehension of the dialogues rather than the preferences for one or the other effect in the audio subtitles. As shown in Figure 37, most of the participants, 34, representing 81% of responses, answered that they had understood the dialogues in the same way. From the remaining participants, 6 (14%) answered that the less acted voice with the audible original, a voice-over effect, posed more problems to understand the dialogues, while only 2 people (5%) reported having had more difficulty understanding the dialogues with a dubbing effect, that means, with an acted voice and without the original dialogue. Therefore, it can be argued that there is no identification of difficulty in comprehension. However, it would be easier to follow this path for future research on the delivery of AST.

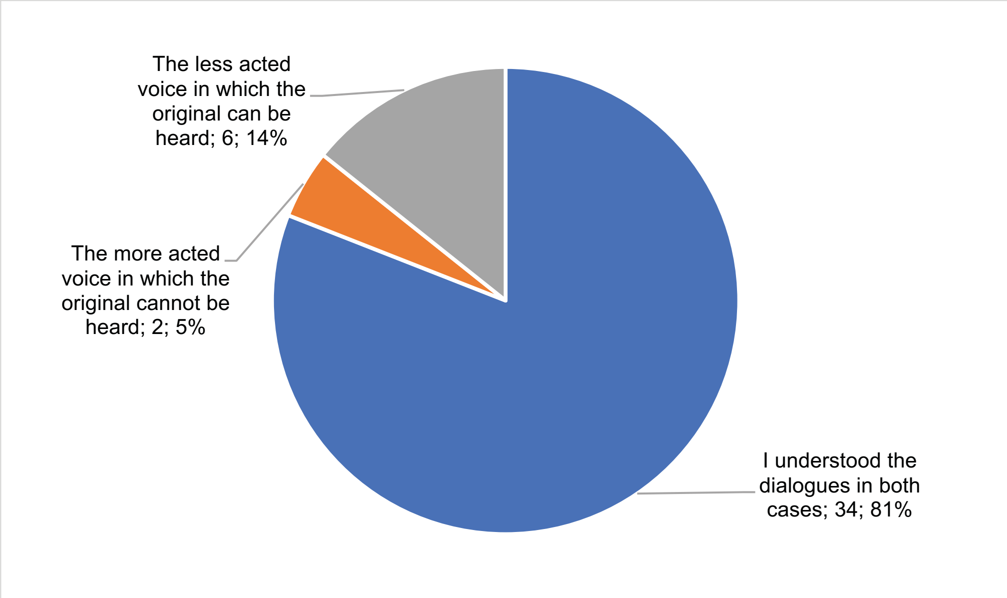


Figure 37. Replies to “Was there an option in which dialogues where easier to understand?”

Which kind of voice do you prefer for the AST (when subtitles are read aloud)?

[¿Qué voz prefieres para los audiosubtítulos (cuando los subtítulos se leen en voz alta)?]

This question centred on the acting of the voices. Within this study, the “acting” was understood as the characteristics found in the prosody and form of dubbing. On the other hand, the less acted voice was the one used in the voice-over effect in which the prosody was closer to the reading-aloud voice.

Having said that, 31 participants (74%) showed a preference for a more acted voice as shown in Figure 38, which would correspond to the prosodic features of dubbing. Seven people (17%) showed a preference for a less acted voice and only 4 (9%) did not show any difference for one or the other effect.

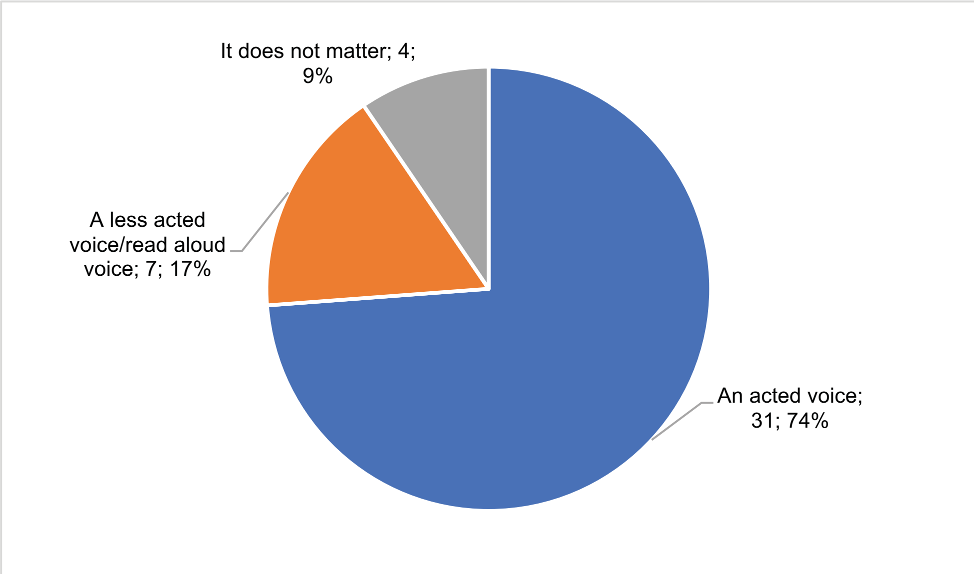


Figure 38. Replies to “Which kind of voice do you prefer for the AST (when subtitles are read aloud)?”

Would you prefer having a different voice for each character?

[¿Preferirías tener una voz diferente para cada personaje?]

The stimuli presented to the participants were treated with only three voice talents: 1 female-voice for all female characters, 1 male voice for all male characters, and a third male voice for the AD. This question was intended to gather information about the preference for different voices for the all the characters in the clips.

In Figure 39, it is shown that a clear majority of people, more accurately 35 respondents (83%), answered that they would like to have a different voice for each character. Of the remaining participants, 5 (12%) did not show a clear opinion arguing that it was not that relevant, and only 2 (5%) said that they did not want different voices.

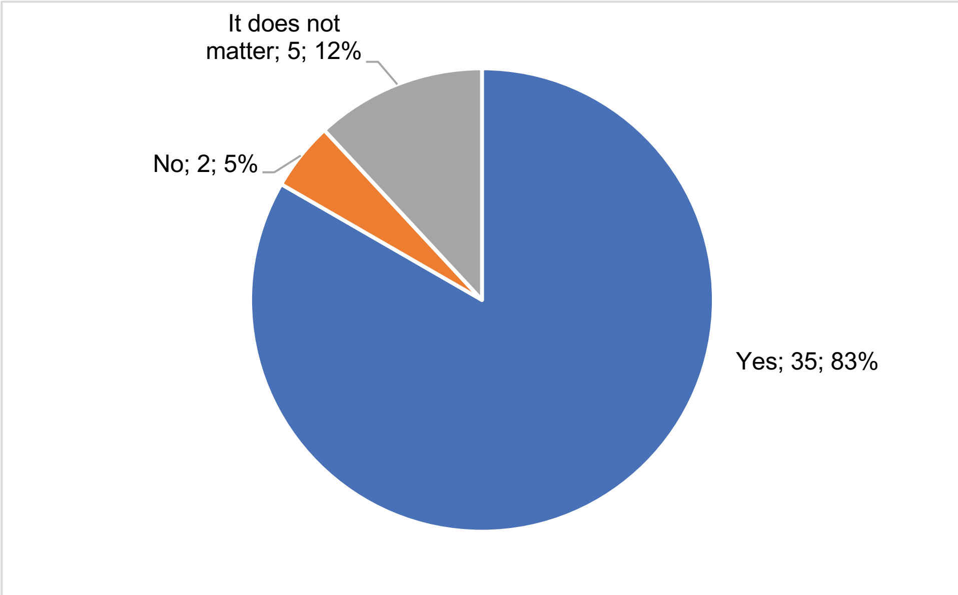


Figure 39. Replies to “Would you prefer having a different voice for each character?”

1.3.2. Audio Description

Although AD was not the main object of study in the experiments, it did play an important role in the reception and comprehension of the clips. As a matter of fact, it is another of the audio inputs received by the participants when exposed to AST. This is why some additional input was gathered on this topic.

Would you like to have a more acted audio description?

[¿Te gustaría tener una audiodescripción más interpretada?]

The question aimed to collect some answers on the characteristics of the voice used in AD. In the stimuli, the same male voice was used and the instructions provided to the professionals were that it should be a neutral voice, with the characteristics found in commercial AD.

As shown in the figure (Figure 40), most part of the respondents (23.55%) answered that they did not want a more acted voice and that they liked the AD provided throughout the clips. Fifteen participants (36%) would rather have a more acted voice since the voice provided was too monotonous for them. Only 4 participants (9%) did not find this point relevant.

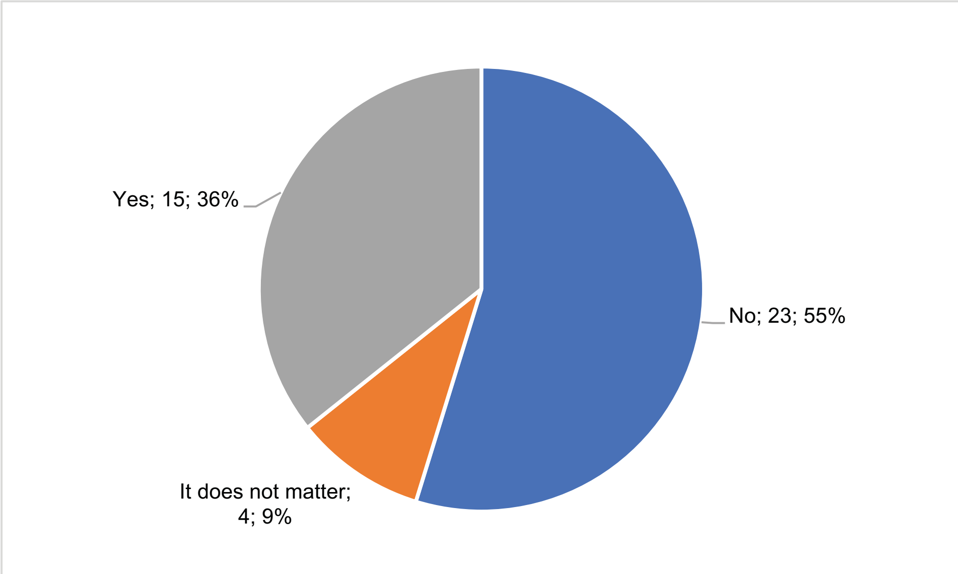


Figure 40. Replies to “Would you like to have a more acted audio description?”

2. Experiment 2

In this section, the results for Experiment 2 are presented and discussed. The presentation and discussion follows the general order: first, results from self-reports instruments is presented, followed by the results obtained from the psychophysiological measurements.

2.1. Results for the Self-Report Questionnaire

The T-SAM is divided into two dimensions: valence and arousal (see Chapter 3, where emotions and their dimensions are discussed). Both dimensions were rated by means of a 9-point scale. For valence, which provides information about the emotion being positive or negative, 1 meant very negative and 9 very positive. For arousal, which provides information about the activation (arousal) caused by the emotion, 1 meant that the person was unaltered completely and 9, that the person was very excited or activated. The results are shown in average marks for each of the dimensions.

2.1.1. Valence

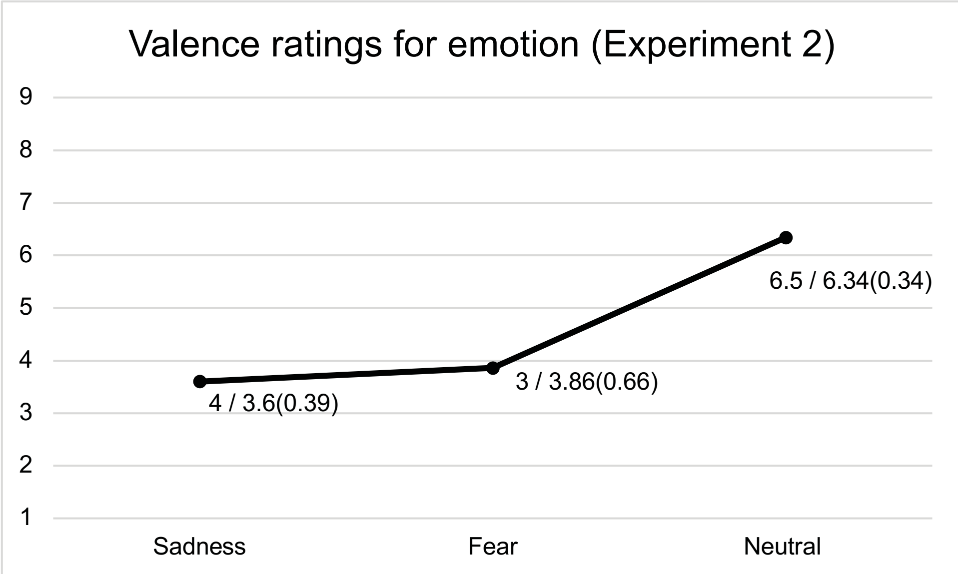


Figure 41. Valence ratings for emotion (Experiment 2): median, average(standard deviation)

Figure 41 shows the median and average results for the valence dimension in Experiment 2. The vertical axis shows the rating from 1 to 9, the horizontal axis shows the three emotional clips (for the emotion of sadness, fear and neutral, correspondingly). In the case of Experiment 2, only one condition was presented to the participants.

A Related-Samples Friedman's Two-Way Analysis of Variance by Ranks was conducted to test the differences among the ratings. The test yielded significant differences (Friedman's $Q_{(2)} = 24.7$; $p < .001$, N=42).

Pairwise comparisons indicated that the ratings for the clips eliciting sadness and fear do not differ significantly (Friedman's $Q_{(1)} = 1.581$; $p = .21$), whereas both ratings are different for the ratings of the clips aimed at a neutral emotion (Sadness/Neutral: Friedman's $Q_{(1)} = 25.0$; $p < .001$; Fear/Neutral: Friedman's $Q_{(1)} = 9.529$; $p = .002$).

The comparison of the averages of the ratings for the dimension of valence provided the results expected according to the literature on the characteristics of emotions (see Chapter 3). Similar to the results acquired in Experiment 1 for the same dimension, the clips portraying sadness (3.6) and fear (3.86) were rated with lower valences in comparison to the clip aimed at a neutral emotion (6.34). The statistical analysis confirms that the differences observed are statistically significant. Both emotions, fear and sadness, are considered as having a low valence according to the literature and the results show that this negativity was correctly induced and perceived by the participants.

2.1.2. Arousal

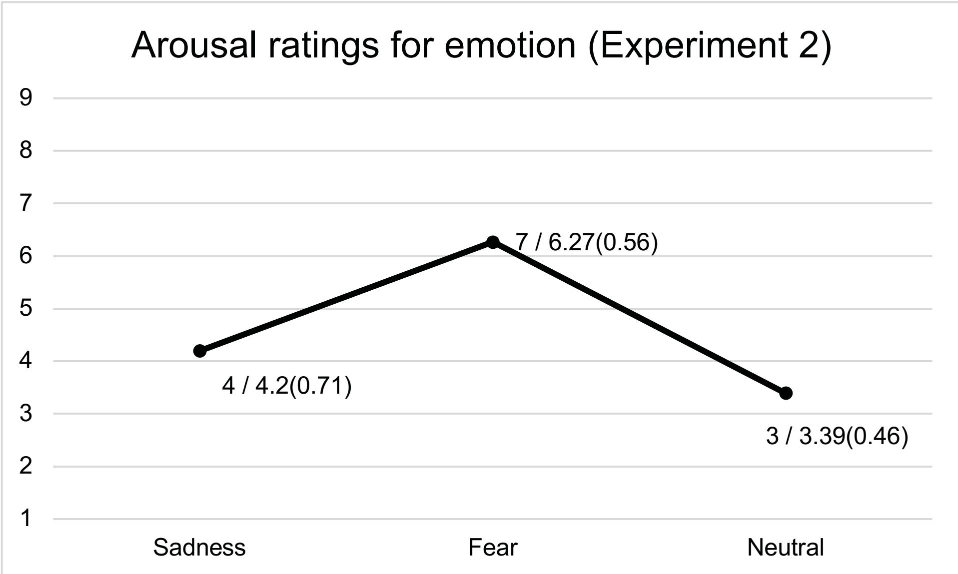


Figure 42. Arousal ratings for emotion (Experiment 2): median, average(standard deviation)

Figure 42 shows the results for the second dimension of the T-SAM questionnaire, arousal. As in the Figure 41, the figure presents only the averages for the condition of emotion: sadness, fear, and neutral.

The Related-Samples Friedman's Two-Way Analysis of Variance by Ranks confirms that the observed differences are significant between the three ratings (Friedman's $Q_{(2)} = 41.367$; $p < .001$). Pairwise comparisons indicate that ratings for the clips portraying sadness and neutral emotion are not statistically significant (Friedman's $Q_{(1)} = 3.789$; $p = .052$); but clips eliciting fear and neutral emotions ratings do yield significance (Friedman's $Q_{(1)} = 21.900$; $p < .001$) as well as the ratings for fear and sadness (Friedman's $Q_{(1)} = 23.684$; $p < .001$).

As expected, the clips inducing emotions of sadness (4.2) and fear (6.27) obtained higher ratings in the dimension of arousal in comparison to the clips portraying a neutral emotion (3.39). The ratings for emotion of fear indicate that the participants were more activated than for the neutral and sadness**.** This considerably relates to the literature and results are similar to those obtained in Experiment 1. Overall, emotions were correctly induced, the ratings provide cohesion between the two experiments and are coherent as far as the literature is concerned (see Chapter 3).

2.2. Results for the Psychophysiological Measures

In this section, the results obtained for the measurements of EDA and HR are presented. The results, expressed in graphs, are accompanied by the statistical analysis.

2.2.1. Results for EDA

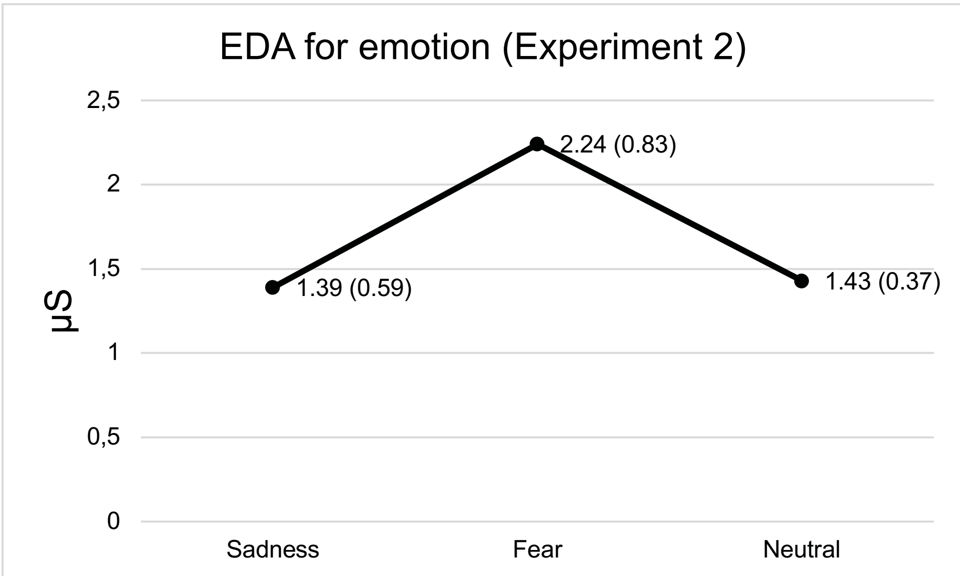


Figure 43. EDA stimuli averages minus 3-minute baseline average for emotion (Experiment 2)

Figure 43 shows the values for the type of emotion. As with Experiment 1 data, a General Linear model of Repeated Measures was used to analyse the possible differences between values of EDA in each condition. The overall comparison does not show differences among the values for each of the emotion-eliciting clips (F (2,70) = 2.110, p = .15, Partial Eta Squared = .057) but pairwise comparisons yield significant differences between those values corresponding to the clips aimed to elicit sadness and fear EDA's ($t_{(35)}$ = -2.717, p = .01).

Significant differences can be found only between the values obtained during the clips aiming for sadness (1.39 μS) and fear emotions (2.24 μS). According to the results, fear-inducing clips result in more activation which is remarkably higher than during sadness-inducing and neutral clips. This pattern correlates to the characteristics of the physiological changes for the emotions which pointed out the general tendency for sadness to decrease EDA.

2.2.2. Results for HR

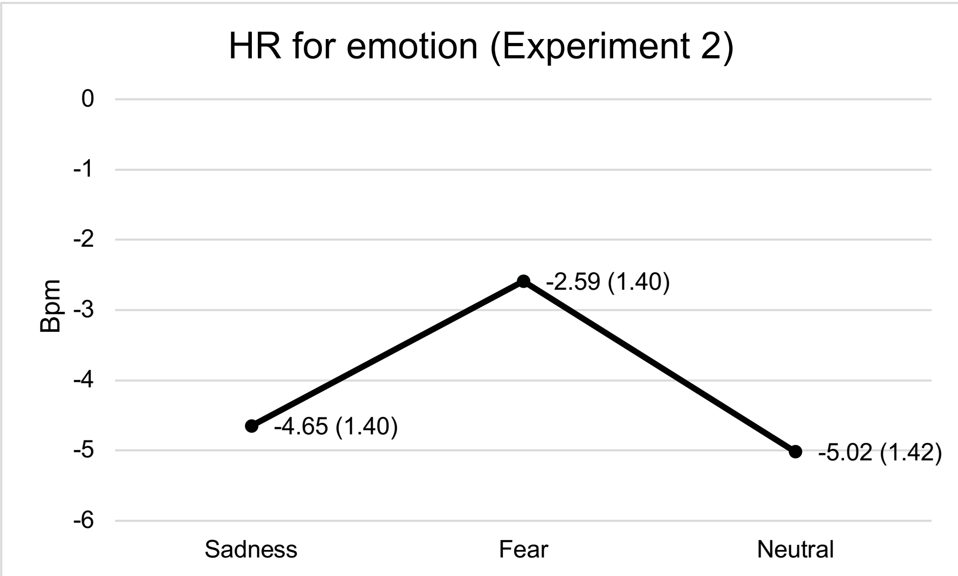


Figure 44. HR stimuli average minus 3-minute baseline average for emotion (Experiment 2)

Figure 44 shows the values for the variable of emotion. The General Linear Model of Repeated Measures applied to the values of the recording of HR does not show any significant difference among the different emotions (F (2, 29) = 2.81, p = .75, Partial Eta Squared = .009).

However, by observing the values it can be highlighted that they are all negative: -5.02 bpm for the neutral emotion-inducing clip, -4.65 bpm for the sadness-inducing clip, and -2.59 bpm for the fear-inducing clip. These results have proven to be inconclusive. Less deceleration is recorded for fear. By taking the explanation that the cognitive process of concentration results in deceleration of the heart (Graham, 1992; Petrie Thomas et al., 2012; Tremayne & Barry, 2001), the neutral-inducing clip and the sadness-inducing clip would result in the participant being more focused in the stimuli.

3. Comparison of results in Experiments 1 and 2

In this section, the results of Experiment 1 and 2 are compared. The comparison will start with the results obtained in the T-SAM questionnaire and then moves onto the results obtained with the recording of EDA and HR.

3.1. T-SAM Questionnaire

The comparison between the valence ratings of blind and partially sighted participants and sighted participants are shown in Figure 45. The analysis reveals significant differences between the ratings of both groups (Mann-Whitney U= 1194, *p* = .004). In general, emotions were induced coherently and similarly in both groups. The ratings were statistically compared by running Mann Whitney tests for each emotion. There are no significant differences between the two groups for the ratings in the clips inducing sadness (Mann-Whitney U = 829.5, *p* = .76) and fear (Mann-Whitney U = 744, *p* = .27). The difference is significant for the ratings of the clip inducing a neutral emotion (Mann-Whitney U = 566, *p* = .006). This shows that sighted participants have rated the clip inducing a neutral emotion as more positive than blind and partially sighted participants. The explanation could be found in the information provided by the visuals or the AD, that might have had an impact on the way the two groups perceived the clip. The colours and general environmental visuals were luminous and colourful while the AD, as stated before, was maintained very neutral, with the intention of not implying any emotion.

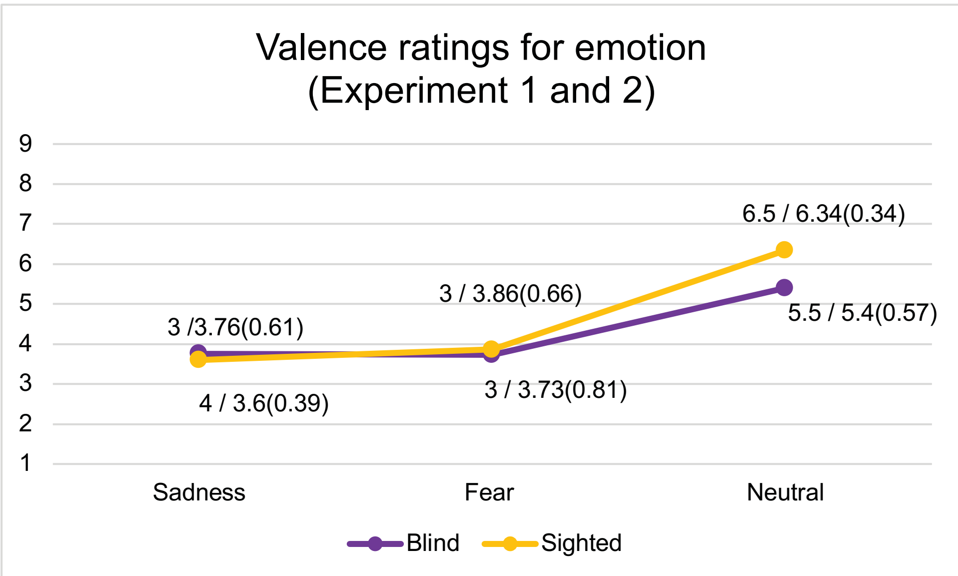


Figure 45. Valence ratings for emotion (Experiment 1 and 2): median, average(standard deviation)

Figure 46 shows the averages of the ratings for the dimension of arousal of both blind, partially sighted and sighted participants. In the case of the ratings of arousal no significant differences were found between the two groups after running Mann-Whitney tests: for clips inducing sadness (Mann-Whitney U = 690, p = .11), fear (Mann-Whitney U = 677.5, p = .08) and a neutral emotion (Mann-Whitney U = 678.5, p = .09). According to the results in the questionnaire, the emotions were induced coherently and similarly in both groups. Emotions were perceived similarly in the clips regardless of the treatment of the clips used to convey the translation of the original language: AST or written subtitles.

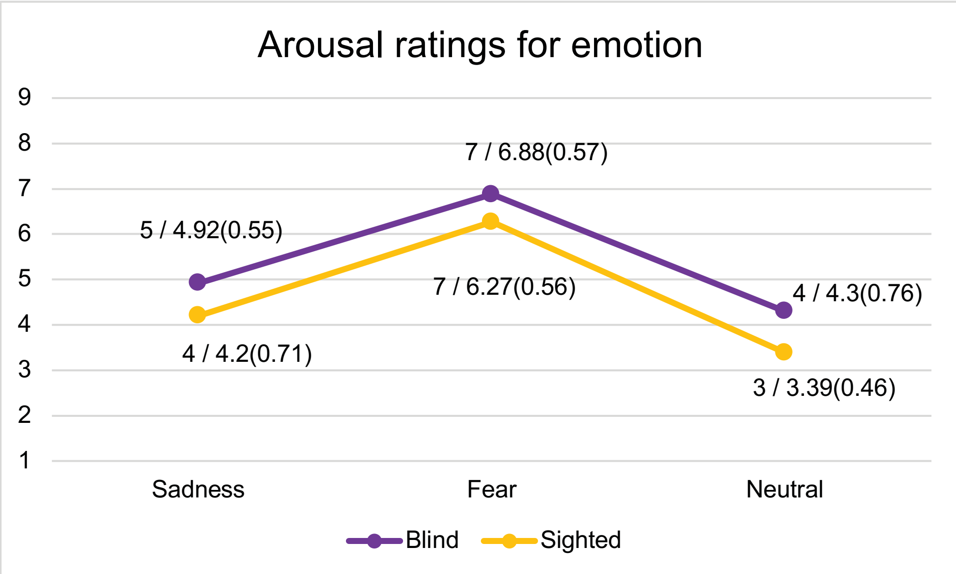


Figure 46. Arousal ratings for emotion (Experiment 1 and 2): median, average(standard deviation)

3.2. Psychophysiological Measures

In Figure 47, the differences between the average values of the stimulus and the averages of the baseline for EDA in blind and partially sighted participants and sighted participants are shown. Comparisons were performed using a T-test for independent samples. Significant differences were found for the values obtained during the clips portraying a neutral emotion ($t_{(67)}$ = -2.512, p = .014) and fear ($t_{(67)}$ = -2,108, p = .039).

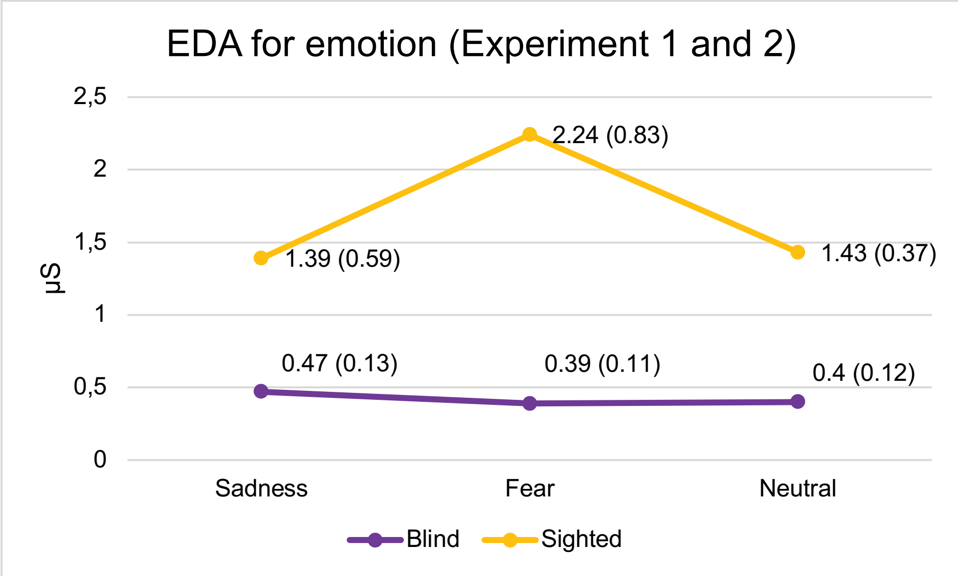


Figure 47. EDA stimuli averages minus 3-minute baseline average for emotion (Experiment 1 and 2)

From the recordings of EDA during the exposure to the clips, it can be observed that sighted participants experienced the clips with more differences among the emotions in comparison to blind and partially sighted participants, in which the values are very similar across the three emotion-inducing clips. It would be worth conducting further research on the impact on visual cues when triggering emotional activation. Again, the fact that the AD and the AST narration aimed to be very plain could provide an explanation for the issue.

The results in sighted participants seem to show a general trend that corresponds to the physiological changes described for each of the emotions since fear has caused higher EDA activation than the other emotions. However, the only significant differences between the groups are found for the values obtained during neutral emotion clips (sighted participants experienced higher arousal than blind and partially sighted participants). Again, the explanation is likely due to the presence of the visuals or the role played by a neutral AD narration.

Regarding HR, in Figure 48, the result of the differences between HR stimuli averages and baseline averages are shown for both groups in Experiments 1 and 2. The results were compared by means of T-tests for independent samples. The tests were performed for the three emotional conditions, sadness, neutral and fear, and did not show any significant difference for any of them.

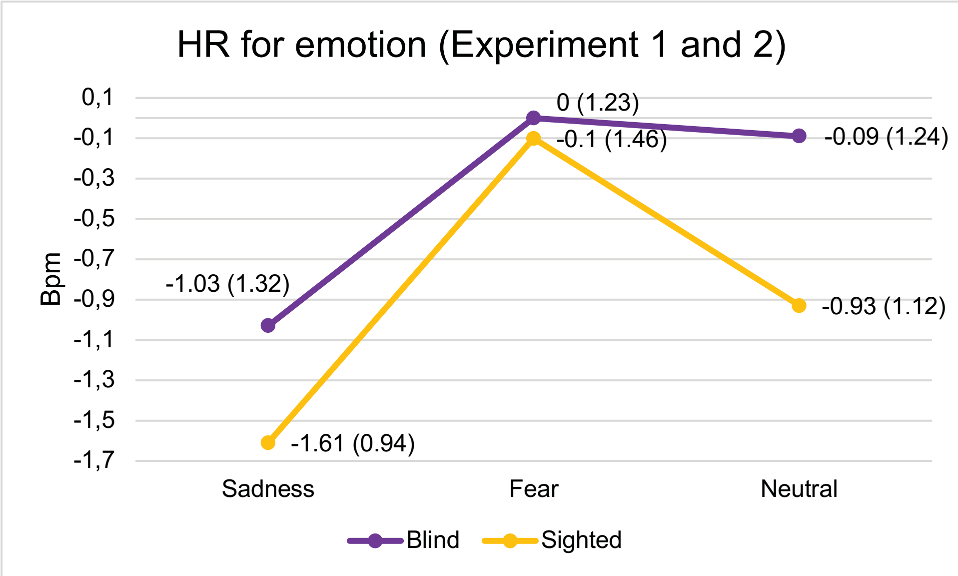


Figure 48. HR stimuli averages minus 3-minute baseline average (Experiments 1 and 2)

Both groups show the same pattern when observing the values. Sadness-inducing clips result in the highest rate of deceleration of the heart, followed by neutral emotion clips and finally fear clips, that cause the lowest rate of deceleration. Sadness and neutral clips would have resulted in more concentration according to the theories that concentration provokes heart rate deceleration (Graham, 1992; Petrie Thomas et al., 2012; Tremayne & Barry, 2001). According to the trends observed, emotions would have been better induced in sighted participants than in blind and partially sighted participants. However, the similarity of the patterns may imply a similar reception of the emotion-eliciting clips for both groups accessing the translation of the audiovisual content through subtitling and audio subtitling, respectively.

4. Correlations

This sections presents the results of the tests runs to assess the relationship between some of the results obtained. In the first part, the results obtained from the self-report instrument in the dimension of valence, that assesses emotional activation, are related to the two psychophysiological measures, EDA and HR. In the second part, the results obtained in the two psychophysiological measures, EDA and HR, are related.

4.1. Self-Report and Psychophysiological Results

The results obtained through the T-SAM questionnaire, for the dimension of arousal, which would subjectively define the emotional activation of the participants, were correlated to the results obtained by means of the recordings of EDA and HR. In order to do so, the ratings in the second scale of the questionnaire (dimension of arousal) were correlated with the differences of the stimuli recording minus the average of the baseline for the two measures.

4.1.1. Correlations in Experiment 1

In order to test the relationship between the results obtained by means of the self-report instrument and psychophysiological measures a Spearman Rank Correlation was run. The relations are made between one of the physiological measures and the ratings of arousal for each of type of stimuli presented to the users, that is, the clip portraying sadness, fear and a neutral emotion. All data seem to show a negative correlation, as presented in the following lines.

For the clips portraying the emotion of sadness, the relationship of EDA results and the ratings of arousal through T-SAM showed very low correlation ($r_s(31) = .14$, $p = .42$). And the same happens between HR results and the ratings of arousal ($r_s(31) = .01$, $p = .91$).

For the clips portraying the emotion of fear, the relationship of EDA results and the rating of arousal showed again very low correlation ($r_s(31) = -.10$, $p = .57$). And the same situation is found between HR results and arousal ratings ($r_s(31) = -.32$, $p = .06$).

For the clips eliciting a neutral emotion, the relationship of EDA results and arousal ratings showed a very low correlation as well ($r_s(31) = .10$, $p = .57$). And it is the same for the correlation between HR results and arousal ratings ($r_s(31) = .03$, $p = .85$).

A Spearman Rank Correlation was also run to assess the relationship of the results obtained from psychophysiological measures and the results in the T-SAM questionnaire for the two effects compared in the experiment.

For the effect of dubbing, results obtained through the recording of EDA and HR were related to the ratings in the arousal dimension of the T-SAM questionnaire. No correlations were found for EDA: sadness ($r_s(17) = .43$, $p =$

.06), fear (r_s (16) = .20, p = .42) and neutral emotion (r_s (14) = .09, p = .71), nor for HR: sadness (r_s (16) = -.06, p = .79), fear (r_s (14) = -.46, p = .06) and neutral emotion (r_s (15) = .01, p = .94).

For the effect of voice-over, results obtained through the recording of EDA and HR were related to the ratings in the arousal dimension of the T-SAM questionnaire. No correlations were found for EDA: sadness (r_s (12) = -.36, p = .20), fear (r_s (13) = -.33, p = .22) and neutral emotion (r_s (15) = .12, p = .64), nor for HR: sadness (r_s (13) = .18, p = .51), fear (r_s (15) = -.05, p = .84) and neutral emotion (r_s (14) = .07, p = .78).

4.1.2. Correlations in Experiment 2

As in Experiment 1, correlations were performed by calculating the Spearman Rank Correlation Coefficient applied to nonparametric measures. The correlations are made between each one of the physiological measures and the ratings in the arousal scale of the T-SAM questionnaire for each of the clips portraying a different emotion: sadness, fear and a neutral emotion. All data seem to show very little correlation, as presented here.

For the clips portraying the emotion of sadness, the correlation of EDA results and the ratings of arousal showed very low results (r_s (34) = -.24, p = .15). And the same goes between HR and the ratings of arousal (r_s (28) = .11, p = .56).

For the clips eliciting the emotion of fear, the correlation between EDA results and arousal ratings in the T-SAM questionnaire showed very low values (r_s (34) = -.19, p = .26). And the same is true for the correlation between HR and arousal ratings (r_s (28) = .09, p = .60).

For the clips portraying a neutral emotion the correlation between EDA and arousal also showed low values (r_s (34) = .02, p = .87) as well as the correlation between HR and arousal ratings (r_s (28) = -.007, p = .97).

In general, correlation has not shown any statistic similitudes between psychophysiological measures and the results obtained in the self-report instrument for any of the emotions in any of the groups. This outcome is similar in other studies within MA using a combination of psychophysiological and self-report methods. In Fryer (2013), scarce isolated correlation was found between the measures. The measure of HR showed certain degree of correlation with certain parts of a questionnaire on presence (confusion, ecological validity and negative effects) only for one of the stimuli, and EDA with only a part of the questionnaire (Identification) for one of the stimuli. In Ramos (2015), HR reported unexpected results as it did not match the results obtained by means of the self-report instrument.

4.2. Psychophysiological Results: EDA and HR

To assess the relationship between the results obtained through the recordings of EDA and HR, Spearman Rank Correlations were conducted. Little correlation was found between all the measures related.

4.2.1. Correlations in Experiment 1

No correlations were found between EDA and HR in the three emotions induced through the stimuli: sadness ($r_s(24) = .17$, $p = .38$), fear ($r_s(24) = .08$, $p = .69$) and neutral emotion ($r_s(24) = .18$, $p = .37$).

Spearman Rank Correlations were also run for each of the two effects. In the clips presenting audio subtitles with a dubbing effect the statistic results did not report significant correlations in any of the three emotions: sadness ($r_s(15) = .14$, $p = .57$), fear ($r_s(11) = -.04$, $p = .88$) and neutral emotion ($r_s(11) = .22$, $p = .45$). The results in the clips showing a voice-over effect are similar: sadness ($r_s(6) = .42$, $p = .28$), fear ($r_s(11) = .19$, $p = .51$) and neutral emotion ($r_s(11) = .03$, $p = .91$).

4.2.2. Correlations in Experiment 2

In the case of Experiment 2, the relationship of results was made only between the three emotions. After running Spearman Rank Correlations, no relevant similitude was found between EDA and HR results in any of the emotions: sadness ($r_s(24) = -.18$, $p = .35$), fear ($r_s(24) = -.18$, $p = .37$) and neutral emotion ($r_s(24) = -.18$, $p = .35$).

Chapter 8

Conclusions

Chapter 8
Conclusions

This thesis had the main aim of comparing the impact of two different strategies for the voicing of audio subtitles (dubbing and voice-over effect) on the emotional activation of blind and partially sighted users by means of different methods and the study of preferences. As secondary objectives, the thesis aimed, first of all, to define and describe AST by means of a literature review and a corpus descriptive study. Secondly, it aimed to present a theoretical framework, focusing on the AST service and on the application of emotional activation as an indicator of user experience. As a third secondary objective, the results obtained from the study of emotional activation when exposing blind and visually impaired participants to audio subtitled stimuli had to be compared to the emotional activation triggered by the same stimuli in their subtitled form to sighted audiences. This chapter presents the main findings of the thesis and discusses its limitations and future research possibilities.

1. Voice-over Effect vs. Dubbing Effect: Emotional Activation and Preferences

To fulfil the thesis' main objective, Experiment 1 in Study 2 was performed. Experiment 1 was conducted with 42 blind and partially sighted participants and data were collected by means of the combination of psychophysiological measures and self-report questionnaires on emotional activation. Furthermore, it also studied the preferences of blind and partially sighted participants in relation to the two AST effects presented. The first hypothesis posed in relation to this aim is the following:

H1. AST with dubbing effect will be more emotionally activating than AST with a voice-over effect.

When it comes to the emotional activation measured through self-report measures, results show that there are only significant differences between the dubbing and the voice-over effect in the ratings of valence for the clips inducing fear. However, by observing the ratings in the self-report instrument some trends can be observed when looking at the interaction of the emotions and the two AST effects.

In valence, that rates emotions as being more or less negative, clips with a dubbing effect were rated as more negative for the clips portraying fear, being this difference being statistically significant. The fact that fear has been rated as more negative in clips with a dubbing effect, and that dubbing has acquired lower rates in comparison to sadness, could suggest that this is the effect that better induces the desired emotions in terms of valence.

In arousal, that rates emotions in terms of more or less activation, the results show that the ratings acquired by the two AST effects are very similar, only showing a slight difference for clips aiming at a neutral emotion, but it did not prove to be significant. Apparently, participants showed more arousal through the dubbing effect when exposed to the clips portraying a neutral emotion.

Psychophysiological measures do not show significant results between the two effects put to the test in this thesis. In EDA, the average values obtained in the recordings for the two effects are very similar. In HR, even if not significant, the difference between the average values for the two AST effects is more remarkable. In this sense, a dubbing effect would cause more deceleration of the heart. According to some scholars, this could be a characteristic of concentration (Graham, 1992; Petrie Thomas et al., 2012; Tremayne & Barry, 2001).

Therefore, the hypothesis can be considered to be partially confirmed, since significant results were only found in the self-report questionnaire for the emotion of fear in which dubbing was rated with less valence. Results using psychophysiological measures do not allow for full confirmation, although the trend observed in the recordings of HR would point to the possibility that a dubbing effect increases the concentration of the user, thus implying more entertainment. The tests run to assess the relationship between self-report and psychophysiological measures or between the two psychophysiological measures, EDA and HR, did not show any significant correlations for any of the two AST effects, which is a similar outcome to the correlations found in previous research (Fryer, 2013; Ramos, 2015). As suggested by Ramos (2015), the physiological reactions might have been related to processes other than emotional arousal, such as concentration or the experimental environment. As mentioned in the limitations of the study (see following Section 8.5.), stimuli with a stronger emotional load could have led to clearer results in physiological recordings.

Regarding preferences, questions were posed in relation to 1) preference for a more acted voice with the original dialogue being unheard (which would correspond to a dubbing effect) or a less acted voice and an audible original (which would correspond to a voice-over effect), 2) if one of these two effects could better transmit the emotion of the scene, and 3) if, in general, participants wished for an acted voice for the AST service. The hypothesis linked to the study of preferences suggested that participants would prefer a dubbing effect.

H2. Participants will show a preference for AST with a dubbing effect.

Results suggest that the hypothesis can be confirmed. There is a slight preference for the option showing a dubbing effect in which the original dialogues are removed (50%) in comparison to those preferring the other option (45%) or those without a preference (5%). Furthermore, participants also suggested pointing the dubbing effect as the better strategy to transmit emotions (52%) in comparison to those preferring the voice-over effect (43%) or those without preference (5%). These answers relate to the general preference for an acted voice in the AST service (74%).

Other questions were posed in relation to the self-reported comprehension of the dialogues and general preferences for the AD voice. When it comes to comprehension, most part of respondents did not report having any problems when understanding dialogues in any of the effects (81%), although more research is needed on this regard with methodologies targeted specifically at comprehension. Also, it is worth mentioning that, whereas a more acted voice is preferred in the AST service, there is a preference for a less acted voice in the AD service (55%).

2. Understanding Audio Subtitling: Definition and Current Practices

The secondary objectives were the definition and description of the AST service by means of a literature survey and a descriptive study of a corpus of films. The thesis is situated in the field of AVT and most particularly, MA. Some works in the field have previously defined AST. For the purposes of this project, the definition by Reviers & Remael (2015, p. 52) was selected:

> AST can therefore be defined as the aurally rendered and recorded version of the subtitles with a film. This spoken version of the subtitles is mixed with the original sound track. AST is usually read, sometimes acted out, by one or more voice actors. Sometimes it is produced by text-to-speech software. The subtitle text is often delivered almost literally, but it can be rewritten to varying degrees, and in addition, the recording method also varies. Usually, AST is recorded as a form of voice-over, which means that the original dialogues can be heard briefly before the translation starts. Sometimes it is recorded in a semi-dubbed form, which means that the original dialogues are substituted by a form of dubbing that is not necessarily entirely lip-sync, that is, synchronous with the lip movement of the speaker.

It was selected because it embraces not only the aural nature of an audio subtitle, but also the differences in their creation and implementation. It also mentions the two effects that were compared in Study 2.

The dubbing and voice-over effects, that are compared in Experiment 1 through emotional assessment and preference study, had also been presented in some works (Braun & Orero, 2010; ISO/IEC 20071-25, 2017; Remael, 2014). However, their characteristics had not been narrowed down. The proposal in this thesis is the following:

1. **Voice-over** effect: in which the AST is displayed a short time after the original, which can be heard underneath. The voice used can be acted to some extent but its prosodic features are related to reading-aloud.

2. **Dubbing** effect: in which the AST is displayed in synchrony with the original line (isochrony), which cannot be heard. The voice used is acted, in the sense that it replicates the emotive lead in the dialogue, similar to the standardised prefabricated natural orality attributed to dubbing.

These effects are not only narrowed down, they were aslo included in a much more general proposal for the treatment of multilingualism through AST, which was also created based on the adaptation of a representation of multilingualism in translation (Sternberg, 1981). This proposal combined the information provided by the AST service and the AD (that are normally combined in multilingual filmic contents), and distinguished:

1. **Vehicular matching:** which implies leaving the original soundtrack without any other information. That means without the existence of an AST or any information provided in the AD track. The comprehension of the content and the language spoken depends only on the receiver's linguistic knowledge.

2. **Selective reproduction:** which implements AST with a voice-over effect in the target language. This lets the user understand that this is a translation. The foreign language is heard in the background without providing further information on the language spoken.

3. **Selective reproduction + language information:** the AST is provided with a voice-over effect, similar to selective reproduction, and the audio describer provides information on the language spoken.

4. **Verbal transposition:** AST goes beyond the translation by imitating and reproducing some of the patterns of the original foreign language spoken. It could be used with both a dubbing or a voice-over effect, since the reproduction of the features concern phonetic, semantic and syntactic structures found in the foreign language that is being portrayed.

5. **Explicit attribution:** is acquired by providing an AST with a dubbing effect after the language to be spoken is announced. By providing an AST with a dubbing effect, the result is closer to the complete homogenisation of the contents.

6. **Homogenising convention:** is simply based on the complete translation of the original line by means of full dubbing, erasing any traces of the foreign language and creating a monolingual discourse.

In order to obtain information on the practice of the AST service in commercialised contents, this thesis also carried out a descriptive study (Study 1). It was performed with a corpus of six multilingual films. This qualitative analysis was based on the following features: voices, assignation, audibility, prosody, isochrony, effect and textual features. Furthermore, the strategies used for the representation of multilingualism as showed in the adaptation conducted with the AST services were also commented upon. The hypothesis was that a high variability in the way the AST service was implemented was to be found.

H3. A high variability in terms of AST delivery is expected to be found.

The results can be summarised as follows for each of the categories under analysis:

Voices: five out of the six films in the corpus had two main voices for the AST service, a female voice for female characters and a male voice for male characters. The AD was narrated by one of the two. The film *Munich* presented a single female voice for the AST and AD services.

Assignation: in half of the films of the corpus there is direct assignation, these films also have a higher presence of audio subtitles. This may be to avoid repetition of the expressions introducing AST and the familiarity of the user with the presence of audio subtitled scenes.

Audibility: in five out of six films, the original lines can be heard under the audio subtitles. Only in *The Science of Sleep*, the volume was very low and made the original barely unheard.

Prosody: audio subtitles tend to be delivered with certain prosodic features closed to acting. The film *Slumdog Millionaire* and *Munich* showed audio subtitles with prosodic features more related to reading aloud.

Isochrony: audio subtitles were presented before or after (if not synchronised) with the original lines, showing great variability in synchronisation. High variability was found not only across films but also within the same film.

Effect: five out of the six films analysed presented an AST service in a style closer to a voice-over effect. In *The Science of Sleep* the effect was closer to a dubbing effect.

Textual features: the main trend across all films is the verbatim reading. That means that unimportant changes have been observed in this analysis between the contents of the written subtitles and their aural counterparts.

The hypothesis is confirmed in general terms. There is no uniformity across the way the AST service is delivered in the films that compose the corpus. However, it can be argued that not all the features here analysed showed the same kind of trend. A voice-over effect is the most used across the films, and this accounts for the fact that the original is normally heard underneath. Isochrony is not full and presents a lot of variability not only across films but also in the same film. When it comes to textual features or the voices used, there is more conformity: audio subtitles are read verbatim and most part of films use two voices (female and male).

The adaptation of the representation of multilingualism to the AST service was created as a theory to highlight the possibilities of this service when providing translation of secondary languages in multilingual films for people who cannot access written subtitles. From the descriptive analysis of films in Study 1, it can be stated that in the corpus most of audio subtitles represent multilingualism by means of the strategy of “selective reproduction” by which a voice-over effect is used and no further information about the linguistic exchange in the scene is mentioned in the AD content. In some cases, “selective reproduction + language information” states the language being translated.

3. Defining a Theoretical Framework: Audiovisual Translation, Media Accessibility and User Experience

Another secondary objective was the creation of a theoretical framework on the experience of emotions through the entertainment of audiovisual contents. The methodology in Study 2 finds its rationale in this theoretical framework. Entertainment is derived from the consumption of audiovisual media. According to the theoretical framework proposed in this thesis, the experience of entertainment is developed by a sum and link of psychological processes. Firstly, entertainment occurs from the experience of flow (Csikszentmihalyi, 1990), and particularly media flow (Sherry, 2004), that results in the concentration into one task. Secondly, being immersed in a product brings to the experience of presence (Lombard & Ditton, 1997), also called transportation, i.e. feeling transported and present in a mediated reality, in this case the audiovisual content. The three concepts have been related to emotional activation and, most particularly, the link between the three of them (entertainment, flow and presence) has been presented in previous studies (Dillon, 2006; Weibel et al., 2008).

The rationale behind the consideration of EDA and HR as psychophysiological measures is found in previous literature that point them out as the most popular measures used in the field (Kreibig et al., 2007) . Both measurements are linked to the activity in the autonomic nervous system and have been used in previous studies. However, a key innovative point in the thesis is applying such measures to the study of the emotional activation triggered by AST. Such measures had been applied in the field of MA in studies on AD (Fryer, 2013; Ramos, 2015) but the inclusion of dialogues translated to the audience by means of audio subtitles had not been put to the test with such methods.

4. Comparing the Emotional Activation of Different User Profiles

As the final secondary objective, the results obtained in the study of emotional activation through both self-report instruments and psychophysiological measurements were compared to those of sighted participants who were exposed to the same audiovisual stimuli provided with written subtitles. This comparison finds its rationale in the fact that an accessible audiovisual material should result in a comparable entertaining or emotional function when compared to its standard commercial counterpart. The way sighted users who have access to the subtitled material receive this type of contents should be comparable to the way blind and partially sighted participants receive the same audio subtitled and audio described contents. The hypothesis suggested that both groups of users, blind, partially sighted, and sighted, will present similar results in emotional activation in both self-report measures and psychophysiological measures.

H4. Similar levels of emotional activation are expected to be found both in sighted audiences consuming subtitles and blind and partially sighted audiences consuming AST without taking into consideration the effect.

From the data obtained in the T-SAM questionnaire, results are very similar between the two groups in both valence and arousal. Furthermore, emotions were correctly induced according to the ratings of valence and arousal, which are concurrent with the literature presented on emotions (see Chapter 3). The emotions of sadness and fear were rated as more negative than the neutral emotion, in the valence scale. Similarly, fear was presented as the emotion reporting higher ratings in arousal, followed by sadness and neutral emotion.

Blind and partially sighted participants and sighted participants reported emotions in a similar way, highlighting the usefulness of the AST service. Participants with no or partial access to written subtitles experienced the audiovisual contents in the same way sighted participants, consumers of the contents in what is normally the standard form of the translation of multilingual contents, i.e. with written subtitles.

Psychophysiological measures did not provide as conclusive results as those obtained through the self-report instrument. In the case of EDA, only differences were found between the two groups for the clips inducing fear and a neutral emotion. In both cases, sighted participants have been proven to show more EDA activation. According to the values obtained, in sighted participants EDA shows that emotions were better induced, as the clips portraying fear show the highest values of activation in comparison to those clips aimed to induce a neutral emotion and sadness.

As far as HR is concerned, none of the statistical analysis yielded significant differences neither between the two groups nor across the three emotions. It is necessary to state that the values obtained through the subtraction of the average of the stimuli minus the average of the baseline are negative. One of the explanations behind this fact is concentration, that is considered by some scholars as resulting in the deceleration of HR (Graham, 1992; Petrie Thomas et al., 2012; Tremayne & Barry, 2001). According to this theory, the clips portraying sadness would have resulted in higher levels of concentration in both groups, followed by those aimed to induce a neutral emotion or fear. However, the pattern shown in the values is similar for both groups, and this provides a well-based starting point for further research on the measure of emotional activation through audiovisual stimuli treated with different access services.

Therefore, the hypothesis was confirmed only partially, since the tests run to assess the relationship between results obtained in the self-repot instrument and psychophysiological results did not prove a significant degree of correlation. Therefore, the results obtained from the self-report instrument, T-SAM questionnaire, confirm that both groups exposed to different treatments of the same audiovisual stimuli (audio subtitled and subtitled) showed similar emotional activation, both in valence and arousal. Furthermore, these results seem to be compatible with the characteristics of the emotions described in Chapter 3. Clips portraying fear and sadness are perceived as more negative than the one aimed at a neutral emotion and fear-inducing clips resulted in higher rates in arousal followed by those portraying sadness and a neutral emotion.

From psychophysiological measures, even if not conclusive, some information can be extracted. EDA proved not to report any significant changes for any of the groups, but it has shown more activation in sighted participants. This could suggest that the visual component improve emotional activation, although previous reception studies on AD comparing the impact of audiovisual contents in blind and partially sighted audiences and sighted audiences suggest that the same experience can be reported regardless of sight condition (Ramos, 2015). The results obtained in HR recordings, even if not significant, show the same pattern of deceleration for both groups across the three emotions, clips portraying sadness resulting in higher level of concentration. The fact that the pattern is similar suggests that clips induced the emotions in a similar way and invites further research with such methodology comparing different services and end users.

5. Main Contributions and Limitations

Other works, such as Master theses (Harrouet, 2016; Thrane, 2013) and other academic productions on integrated AST (Benecke, 2012; Braun & Orero, 2010; Remael, 2012) and computer systems for the delivery of synthetic AST (Mihkla et al., 2014; Nielsen & Bothe, 2008; Verboom et al., 2002) have touched upon the aural delivery of written subtitles for those audiences without no or partial access to the visual component in audiovisual contents. This is the first PhD thesis devoted solely to the study of the AST service to the best of the author's knowledge.

This thesis has had the aim to collect all the literature written up to date about AST. The contents have shown to not only be wide but also interdisciplinary, drawing information from fields such as AVT and MA, and telecommunications engineering. The state of the art includes a revision of regulations, standards and guidelines taking into consideration only the AST service, something that had only been done before on media access services in general.

In Chapter 2, two main contributions have been presented. First, the two strategies previously mentioned for the voicing of subtitles, referred to as "effects", dubbing and voice-over effect, have been well defined and their features described. Second, by taking into consideration the definition of these two effects, a proposal for the treatment of multilingualism through the AST service has been addressed based on a representation of multilingualism in text translation (Sternberg, 1981), which had been previously used for SDH (Szarkowska et al., 2013).

Another of the contributions of this thesis is developing a theoretical framework on the experience of emotions derived from the entertainment experienced during the consumption of audiovisual contents. The theoretical framework has described how this emotional activation can be measured by different methods. Previous works have studied emotional activation when exposing participants to audiovisual stimuli but it has seldom been applied in the field of media accessibility. Most particularly, the experiments in this thesis apply for methods on emotional activation for the first time, namely the measurement of EDA and HR and a questionnaire on emotional activation, to the study of AST reception and different AST strategies. The emotional activation elicited by the experience of entertainment provides support to the application of a methodology that combines self-report measures, deemed to be subjective, with psychophysiological measures, deemed to be objective. Such methodologies can be further exploited in this line of research.

When it comes to self-report questionnaires, a contribution of this thesis is the adaptation of the SAM questionnaire, based on graphic and numeric scales, to the needs of people who are blind or partially sighted. Although it is true that other scholars had proposed adaptations, such as sliding Likert-scales (Fryer, 2013), the T-SAM, which is based on graphic elements, has proven to be a tool useful for all kinds of visual conditions. Not only for its haptic use, but also because it has simplified and augmented all the contents and has incorporated Braille and augmented numbers, that can be used for both partially sighted and sighted persons.

However, during the development of the project some limitations have been encountered. One of the limitations of this project was the collection of all documents presenting information on the AST service. Although the best of efforts were made to maintain the relevant information up to date, a large list of countries were considered. Different laws and acts are passed continuously in the countries included in the literature review. Furthermore, such legal documents are found in the language of the country or territory and the English translations can take some time. In the case of European countries, such state regulations will see amendments and changes in the next three years as indicated in the contents found in the new directive that is going to be passed by the European Parliament very soon.

The descriptive analysis of audio subtitled commercialised films found in this thesis was conducted with a corpus of 6 films. The access services of these films were created in the UK and the United States only. Multilingual films are found around the globe and AST services are provided by industries in different countries. A small-scale qualitative analysis was prioritized, as it was not the central element of the thesis and it only aimed to support the literature review and stimuli creation for the experiments. A larger study would be needed to shed light on the specificities of AST across countries, languages and content types.

When it comes to the experiments, Experiment 1 found in this thesis was conducted with 42 participants who are blind or partially sighted. The number makes a very decent sample in studies of reception within the field of MA, in which is recommended to have a minimum of 30 participants having in mind a number of 5 of which no complete data will be acquired (Orero et al., 2018). However, for psychophysiological measures a larger sample can provide more clarity to the results obtained.

The recording of psychophysiological measurements has proven to be a very complex task. The environment in which such recordings are conducted should be extremely controlled, as well as the characteristics of the stimuli used. Although all the experiments were performed in the same room, it was not a specific lab in which all factors could be full controlled.

Despite the limitations, the thesis has come up with multiple contributions both at a theoretical and practical level. The study of emotional activation proposes methodologies that have not been introduced fully in the field of MA and in testing the experience of end users. Even if psychophysiological measurements have not reported statistically significant results, some trends have been observed and the refinement of their application could lead to promising results in combination with other tools, such as self-report instruments. This thesis opens up new paths to follow in the study of the experience of end users of media access services.

Further research could be conducted on the adaptation of instruments for the emotional assessment of persons who are blind and partially sighted. Furthermore, this type of questionnaire on emotional assessment can be applied to a variety of studies, adapting them to people with specific needs could improve autonomy of the participants and increase the reach of the studies.

The use of psychophysiological recordings with emotion-eliciting audiovisual materials has been done in the past. Such studies have validated scenes which portray deaths or other extremely dramatic events. This study made use of scenes of dialogues, which is a type of stimuli which has not been used in MA up to date with such methodology. However, scenes portraying extreme emotions with dialogues are difficult to find. In future studies, the use of clips portraying more extreme and detailed emotions could lead to emotional reactions easier to record and analyse.

To sum up, the development of this project has provided interesting contributions for the development and evolution of AST. The passing of new regulations and standards and the introduction of new technologies in the area of media access services have highlighted the need and role of the AST service. Thus, this work finds a niche in current research in the field of MA and AVT in general. On the one hand, it contributes a solid literature survey on the AST service, its categorisation and characteristics, a descriptive analysis of audio subtitled contents, and a proposal on the treatment of multilingual contents through AST. On the other hand, it presents and implements a methodology which has seldom been used in MA research and that opens new avenues for future research on user experience through emotional activation.

References

References

Accessible Instructional Materials For Maine Students With Print Disabilities (n.d.). *What is a print disability?*. Augusta: Maine CITE Coordinating Centre on behalf of the Maine Department of Education.

Accessible Media Inc. & The Canadian Association of Broadcasters. (2014). *Described Video Best Practices. Artistic and Technical Guidelines*. Toronto, Canada: Accessible Media Inc.

Ackerman, B. P., Abe, J. A., & Izard, C. E. (1998). Differential emotions theory and emotional development: Mindful of modularity. In M. F. Mascolo & S. Griffin (Eds.), *What Develops in Emotional Eevelopment? Emotions, Personality, and Psychotherapy* (pp. 85–106). Boston, MA: Springer

Act on Croatian Radio-television (2012). Official Gazette 137/10, 76/12, Zagreb, Croatia, 6 July 2012.

Act on Czech Television (1992). Official Gazette 483/1991, Praga, Czech Republic, 1 January 2992.

Act on Electronic Communications (2007). Official Gazette 28/29.03.2018, 29 March 2018, Sofia, Bulgaria.

Act on Radio and Television (2018). Official Gazette Poz. 915, Warsaw, Poland, 22 March 2018.

Act on Television and Radio Operations (2010). Official Gazette 731/2010, Helsinki, Finland.

Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion*, *30*(2), 179–188. Doi: 10.1007/s11031-006-9020-2

Asociación Española de Normalización (2005). *Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías* (UNE 153020:2005). Retrieved from https://www.aenor.com/normas-y-libros/buscador-de-normas/une/?c=N0032787

Albert, W., & Tullis, T. (2013). *Measuring the User Experience : Collecting, Analyzing, and Presenting Usability Metrics* (2nd ed.). Burlington, MA: Morgan Kaufmann.

Aluja, A., Rossier, J., García, L. F., Angleitner, A., Kuhlman, M., & Zuckerman,

M. (2006). A cross-cultural shortened form of the ZKPQ (ZKPQ-50-cc) adapted to English, French, German, and Spanish languages. *Personality and Individual Differences*, *41*(4), 619–628. Doi: 10.1016/j.paid.2006.03.001

American Printing House for the Blind Inc (1997). *Guidelines for Design of Tactile Graphics*. Retrieved from http://www.aph.org/research/guides/ (December 2018)

Audiovisual Media Services Directive (2010). Official Journal of the European Union L 95/1 (Directive 2010/13/EU).

Audiovisual Media Services Directive (2011). Official Gazette 87/2011, Ljubljana, Slovenia.

Baños, R., & Chaume, F. (2009). Prefabricated orality. A challenge in audiovisual translation. In M. G. Marrano, G. Nadini, & C. Rundle (Eds.), [Special Issue] *InTRAlinea: The Translation of Dialects in Multimedia* (n.p.) Retrieved from http://www.intralinea.org/specials/article/Prefabricated_Orality (January 2019)

Baños, R. M., Botella, C., Alcañiz, M., & Liaño, V. (2004). Immersion and emotion: Their impact on the sense of presence. *CyberPsychology & Behaviour*, *7*(6), 734–741.

Barr, W., Hodge, S., Leeven, M., Bowen, L., & Knox, P. (2012). Emotional support and counselling for people with visual impairment: Quantitative findings from a mixed methods pilot study. *Counselling and Psychotherapy Research*, *12*(4), 294–302. Doi: 10.1080/14733145.2012.663776

Bartoll, E. (2006). Subtitling multilingual films. In M. Carroll, H. Gerzymisch-Arbogast, & S. Nauert (Eds.) *MuTra 2006 – Audiovisual Translation Scenarios: Conference Proceedings, 1-5 May, Copenhagen, Denmark* (pp. 1–6).

Bartsch, A., Mangold, R., Viehoff, R., & Vorderer, P. (2006). Emotional gratifications during media use - An integrative approach. *Communications*, *31*(3), 261–278. Doi: 10.1515/COMMUN.2006.018

Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology*, *60*(1), 34–43. Doi: 10.1016/j.ijpsycho.2005.04.007

Benecke, B. (2012). Audio description and audio subtitling in a dubbing country : Case studies. In E. Perego (Ed.), *Emerging Topics in Translation: Audio Description* (pp. 99–104). Trieste, Italy: Edizioni Università di Trieste.

Benecke, B., & Dosch, E. (2004). *Wenn aus Bildern Worte werden. Durch Audio-Description zum Hörfilm*. Munich, Germany: Bayerischer Rundfunk.

Bernat, E., Patrick, C. J., & Benning, S. D. (2008). Effects of picture content and intensity on affectiv ephysiological response. *Psychophysiology*, *43*(1), 93–103. Doi: 10.1111/j.1469-8986.2006.00380.x

Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators and Virtual Environments*, *12*(5), 456–480. Doi: 10.1162/105474603322761270

Biocca, F., & Kim, T. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-Mediated Communication*, *3*(2). Doi: 10.1111/j.1083-6101.1997.tb00073.x

Bleichenbacher, L. (2008). *Multilingualism in the Movies: Hollywood Characters and Their Language Choices.* Tübingen, Germany: Francke Verlag.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Semantic Differential Manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(I), 49–59. Doi: 10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and psychology. In R. D. Lane, & L. Nadel (Eds.), *Series in Affective Science. Cognitive neuroscience of emotion* (pp. 242–276). New York, NY: Oxford University Press.

Braun, S. (2008). Audiodescription research: State of the art and beyond. *Translation Studies in the New Millennium*, *6*, 14-30.

Braun, S., & Orero, P. (2010). Audio description with audio subtitling – an emergent modality of audiovisual localisation. *Perspectives: Studies in Translatology*, *18*(3), 173–188. Doi: 10.1080/0907676X.2010.485687

Bréan, S., & Cornu, J.-F. (2012). The translation and reception of multilingual films. *InMedia (online)*, *2*, 1–5. Retrieved from http://inmedia.revues.org/486 (December 2018)

Broadcasting Act (2015). Broadcasting Authority Cap. 350, Valletta, Malta.

Broadcasting Authority of Ireland (2012). *Guidelines – Audio Descritpion*. Retrieved from https://www.bai.ie/en/codes-standards/ (December 2018)

Broadcasting Act (2018). Official Gazette S.C. 1991, c. 11, 12 December 2018, Ottawa, Canada.

Broadcasting Services Act (2018). Official Gazette No. 110, 1992 (Compliation No. 94), 1 September 2018, Canberra, Australia.

Bryant, J., & Miron, D. (2002). Entertainment as media effect. In J. Bryant, & D. Zillmann (Eds.) *Media Effects* (2nd ed., pp. 549–582). Mahwah, NJ: Erlbaum.

Busarello, E., & Sordo, F. (2011). *Manuale per aspiranti descrittori di audiofilm per non vedenti*. Scurelle, Italy: Cooperativa Sociale Senza Barriere ONLUS

Cabeza-Cáceres, C. (2010). Opera audio description at Barcelona's Liceu theatre. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds.), *New Insights into Audiovisual Translation and Media Accessibility* (pp. 227–237). Amsterdam: Rodopi.

Cabeza-Cáceres, C. (2013). *Efecte de la velocitat de narració, l'entonació i l'explicitació en la comprensió fílmica*. Universitat Autònoma de Barcelona. Retrieved from http://hdl.handle.net/10803/113556 (December 2018)

Cabeza-Cáceres, C., & Matamala, A. (2008). La audiodescripción de ópera: una nueva propuesta. In Á. Pérez-Ugena, & R. Vizcaíno-Laorga (Eds.), *Ulises y la comunidad sorda* (pp. 95–108). San Lorenzo del Escorial, Spain: Observatorio de las Realidades Sociales y de la Comunicación.

Caillé, P.-F. (1960). Cinéma et traduction: le traducteur devant l'écran. *Babel*, *6*(3), 103–109. Doi: 10.1075/babel.6.3.01cai

Cantor, J. (2002). Fright reactions to mass media. In J. Bryant, & D. Zillmann (Eds.), *Media effects: Advances in Theory and Research* (pp. 287–306). Mahwah, NJ: Erlbaum.

Carroll, N. (1990). *The Philosophy of Horror, Or Paradoxes of the Heart*. *Philosophy and Literature* (Vol. 16). New York, NY: Routledge. Doi: 10.1353/phl.1992.0012

Cary, E. (1960). La traduction totale. *Babel*, *6*(3), 110–115. Doi: 10.1075/babel.6.3.02car

Centre for Equitable Library Access. (n.d.). *What is a print disability?* Retrieved from http://iguana.celalibrary.ca/iguana/www.main.cls?surl=CELA-WhatIsAPrintDisability&lang=eng&theme=reset (May 2017)

Centre for Excelence in Universal Design (2013). *Spoken Subtitles*. Retrieved from http://universaldesign.ie/Technology-ICT/Irish-National-IT-Accessibility-Guidelines/Digital-TV-equipment-and-services/guidelines-for-digital-tv-equipment-and-services/Spoken-subtitles/ (June 2017)

Chaume, F. (2004). Synchronization in dubbing. A translational approach. In P. Orero (Ed.), *Topics in Audiovisual Translation* (pp. 35–57). Philadelphia, PA: John Benjamins.

Chaume, F. (2012). *Audiovisual Translation: Dubbing*. Manchester: St. Jerome.

Chmiel, A., & Mazur, I. (2016). Researching preferences of audio description users — Limitations and solutions. *Across Languages and Cultures*, *17*(2), 271–288. Doi: 10.1556/084.2016.17.2.7

Christie, I. C., & Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: A multivariate approach. *International Journal of Psychophysiology*, *51*(2), 143–153. Doi: 10.1016/j.ijpsycho.2003.08.002

Clasen, M. (2017). *Why Horror Seduces*. Oxford, UK: Oxford University Press.

Code of Audivisual Rights (2017). State Agency Official Gazette, 24 November, Madrid, Spain.

Coordinated Decree on Audiovisual Media Services (2018). Superior Council of Audiovisual Media, 21 August 2018, Bruxelles, Belgium.

Corral, A., & Lladó, R. (2011). Opera Multimodal Translation: Audio Describing Karol Szymanowski's Król Roger for the Liceu Theatre, Barcelona. *The Journal of Specialised Translation*, 15, 163–179. Retrieved from http://www.jostrans.org/issue15/art_corral.php (September 2018)

Croatian Radio-television (2013). Contract between Croatian Radio-telvision and the Government of the Republic of Croatia 2013-2017, Zagreb, Croatia.

Csikszentmihaly, M. (1988). The future of Flow. In M. Csikszentmihalyi, & I. S. Csikszentmihalyi (Eds.), *Optimal Experience: Psychological Studies of Flow in Consciousness* (pp. 364–383). New York, NY: Cambridge University Press.

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikszentmihaly, & I. S. Csikszentmihaly (Eds.), *Optimal Experience: Psychological Studies of Flow in Consciousness* (pp. 15–35). New York, NY: Cambridge University Press.

Csikszentmihalyi, M. (1993). *The Evolving Self: A Psychology for the Third Millennium*. New York, NY: Harper Perennial.

Csikszentmihalyi, M. (1997). *Finding Flow: The Psychology of Engagement with Everyday Life*. New York, NY: Basic Books.

Csikszentmihalyi, M., & Czikszentmihalyi, I. S. (1988). *Optimal Experience: Psychological Studies of Flow in Consciousness*. New York, NY: Cambridge University Press.

Culture and Accessibility - Public Radio and Television 2014-2019 (2014). Official Gazette 2012/13:164, 20 June 2013, Stockholm, Sweden.

Daltrozzo, J., Wioland, N., Mutschler, V., Lutun, P., Calon, B., Meyer, A., & Kotchoubey, B. (2010). Emotional electrodermal response in coma and other low-responsive patients. *Neuroscience Letters*, *475*(1), 44–47. Doi: 10.1016/j.neulet.2010.03.043

De Higes-Andino, I. (2014). The translation of multilingual films: Modes, strategies, constraints and manipulation in the Spanish translations of It's a Free World… *Linguistica Antverpiensia. New Series – Themes in Translation Studies*, *13*, 211–231.

de Jong, F. (2004). *Access Services for digital television*. Grand-Saconnex, Switzerland: European Broadcasting Union

Decreet houdende wijziging van diverse bepalingen van het decreet van 27 maart 2009 betreffende radio-omroep en televisie (2012). Flemish Government 2009/35356, 30 April 2009, Brussels, Belgium.

Derbring, S., Ljunglöf, P., & Olsson, M. (2009). SubTTS: Light-Weight automatic reading of subtitles. In K. Jokinen, & E. Bick (Eds.), *Proceedings of the 17th Nordic Conference on Computational Linguistics*, *May 14-19, Odense, Denmark*. 272–274.

Detenber, B., Simons, R., & Bennett, G. (1998). Roll 'em!: The effects of picture motion on emotional responses. *Journal of Broadcasting & Electronic Media*, *42*(1), 113–127. Doi: 10.1080/08838159809364437

Di Giovanni, E. (2014a). Audio introduction meets audio description: An Italian experiment. In R. M. Bollettieri Bosinelli, E. Di Giovanni, & L. Rossato (Eds.), [Special Issue] *inTRAlinea: Across Screens Across Boundaries*. Retrieved from http://www.intralinea.org/specials/article/2072 (September 2018)

Di Giovanni, E. (2014b). Visual and narrative priorities of the blind and non-blind: eye tracking and audio description. *Perspectives: Studies in Translation Theory and Practice*, *22*(1), 136–153. Doi: 10.1080/0907676X.2013.769610

Díaz Cintas, J., & Anderman, G. (2009). The wealth and scope of audiovisual translation. In J. Díaz Cintas, & G. Anderman (Eds.), *Audiovisual Translation: Language Transfer on Screen* (pp. 1–17). Basingstoke, UK: Pelgrave Macmillan.

Díaz Cintas, J., & Orero, P. (2006). Voice-over. In K. Brown (Ed.), *Encyclopedia of Languages & Linguistics* (Vol. 13, pp. 477–479). Oxford, UK: Elsevier.

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., & Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Frontiers in Psychology*, *6*, 1–9. Doi: 10.3389/fpsyg.2015.00026

Dillon, C. (2006). *Emotional Responses to Immersive Media*. London: University of London. Retrieved from https://research.gold.ac.uk/view/goldsmiths/Dillon=3ACatherine=3A=3A.html (December 2018)

Dillon, C., Keogh, E., Freeman, J., & Davidoff, J. (2001). Presence: Is your heart in it. In *Proceedings of the 4th Annual International Workshop on Presence, May 21-23, Philadelphia, Pennsylvania*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.491.5878 (September 2018)

Eardley-Weaver, S. (2014). *Lifting the Curtain on Opera Translation and and Accessibility: Translating Opera for Audiences with Varying Sensory Ability*. Durham, UK: Durham University. Retrieved from http://etheses.dur.ac.uk/10590/ (December 2018)

Ekman, P. (1992). Are there basic emotions?. *Psychological Review*, 99(3), 550-3. Doi: 10.1037/0033-295X.99.3.550

Ekman, P. (2005). Basic emotions. In T. Dalgleish, & M. J. Power (Eds.) *Handbook of Cognition and Emotion* (pp. 45–60). Chichester, UK: John

Wiley & Sons, Ltd. Doi: 10.1002/0470013494.ch3

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. Doi: 10.1177/1754073911410740

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, *53*(4), 712–717. Doi: 10.1037/0022-3514.53.4.712

Electronic Mass Media Law (2010). Official Gazette 118 (4310), 12 July 2010, Riga, Latvia.

Electronic Media Act (2013). Official Gazette 153/09, 84/11, 94/13, 30 July 2013, Zagreb, Croatia.

Entidade Reguladora para a Comunicação Social (2016). Plano pluranual que define o conjunto de obrigações relativas à acessibilidade dos seriços de programas televisivos e dos serviços audiovisuais a pedido por pessoas com necessidades especiais (ERC/2016/250), Lisboa, Portugal.

Equality Act (2010). Parliament of the United Kingdom Chapter 15, 8 April 2010, London, United Kingdom.

Espasa, E., & Corrius, M. (2017). The audiodescription of multilingual films : audiosubtitling and beyond. Paper presented at the *Advanced Research Seminar in Audio Description 2017*, 16-17 March, Barcelona, Spain. Retrieved from http://grupsderecerca.uab.cat/arsad/sites/grupsderecerca.uab.cat.arsad/files/corrius_espasa_arsad_2017.pdf (January 2019)

Etzel, J. A., Johnsen, E. L., Dickerson, J., Tranel, D., & Adolphs, R. (2006). Cardiovascular and respiratory responses during musical mood induction. *International Journal of Psychology*, *61*, 57–69. Doi: 10.1016/j.ijpsycho.2005.10.025

European Platform of Regulatory Authorities (2018). *Media Legislation.* Retrieved from https://www.epra.org/articles/media-legislation (January 2019)

European Union Agency for Fundamental Rights (2014). *Accessibilty standards for audio-visual media. Indicators on political participation of persons with disabilities*. Vienna: European Union Agency for Fundamental Rights.

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. Doi: 10.1037/0022-3514.74.4.967

Fernández-Torné, A. (2016). *Audio Description and Technologies. Study on the Semi-Automatisation of the Translation and Voicing of Audio Descriptions*. Bellaterra, Spain: Universitat Autònoma de Barcelona. Retrieved from https://www.tesisenred.net/handle/10803/394035 (December 2018)

Fernández-Torné, A., & Matamala, A. (2016). Text to speech vs. human voiced audio descriptions: A reception study in films dubbed into Catalan. *The Journal of Specialised Translation*, *24,* 61–68.

Fernández, E. I., Martínez, S. M., & Núñez, A. J. C. (2015). Cross-fertilization between reception studies in audio description and interpreting quality assessment: The role of the describer's voice. In R. Baños, & J. Díaz Cintas (Eds.), *Audiovisual Translation in a Global Context* (pp. 72–95). London, UK: Palgrave Macmillan. Doi: 10.1057/9781137552891_5

Fernández Megías, C., Pascual Mateos, J. C., Soler Ribaudi, J., & Fernández-Abascal, E. G. (2011). Validación española de una batería de películas para inducir emociones. *Psicothema*, *23*(4), 778–785. Retrieved from http://www.psicothema.com/psicothema.asp?ID=3956 (December 2018)

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050–1057. Doi: 10.1111/j.1467-9280.2007.02024.x

Freeman, J. (2003). Converging evidence on the structure of presence. In *6th International Workshop on Presence*, *October 6-8, Aalborg, Denmark*. Retrieved from https://ispr.info/presence-conferences/previous-conferences/presence-2003/ (December 2018)

Freeman, J. (2004). Implications of the measurement of presence from convergent evidence on the structure of presence. In *4th International Conference of the Communication Association, May 21-23, New Orleans, Louisiana.* Retrieved from https://ispr.info/presence-conferences/previous-conferences/presence-at-ica-2004-proceedings/ (December 2018)

Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., & Ijsselsteijn, W. (2000). Using behavioral realism to estimate presence: a study of the utility of postural responses to motion stimuli. *Presence*, *9*(2), 149–164. Doi: 10.1162/105474600566691

Freeman, J., Lessiter, J., Pugh, K., & Keogh, E. (2005). When presence and emotion are related, and when they are not. In Slater, M. (Ed.) *Proceedings of the 8th Annual International Workshop on Presence*, *September 21-23, London, UK*. Retrieved from https://ispr.info/presence-conferences/previous-conferences/presence-2005/ (January 2019)

Fresno Cañada, N. (2014). *La (re)construcción de los personajes fílmicos en la audiodescripción. Efectos de la cantidad de información y de su segmentación en el recuerdo de los receptores.* Bellaterra, Spain: Universitat Autònoma of Barcelona. Retrieved from https://www.tdx.cat/handle/10803/285420 (January 2019)

Frijda, N. H. (1993). Moods, emotion episodes, and emotions. In M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of Emotions* (pp. 381–403). New York, NY: Guilford Press.

Fryer, L. (2013). *Putting It Into Words: The Impact of Visual Impairment on Perception, Experience and Presence*. London, UK: University of London. Retrieved from http://research.gold.ac.uk/view/goldsmiths/Fryer=3ALouise=3A=3A.html (January 2019)

Fryer, L., & Freeman, J. (2013). Visual impairment and presence. In H. Witchel (Ed.), *Proceedings of the 2013 Inputs-Outputs Conference on An Interdisciplinary Conference on Engagement in HCI and Performance - Inputs-Outputs ’13, June 26, Brighton, UK* (pp. 1–5). New York, NY: ACM Press. Doi: 10.1145/2557595.2557599

Fryer, L., & Freeman, J. (2014). Can you feel what I’m saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment. In A. Felnhofer, & O. D. Kothgassner (Eds.), *Proceedings of the International Society for Presence Research 2014, March 17-19, Vienna, Austria* (pp. 99–107).

General Law on Aduiovisual Communication (2015). Official Gazette 7/2010, 1 May 2015, Madrid, Spain.

Gendolla, G. H. E. (2000). On the impact of mood on behavior: An integrative theory and a review. *Review of General Psychology*, *4*(4), 378–408. Doi: 10.1037/1089-2680.4.4.378

Georgakopoulou, Y. (2008). *Audio Description Guidelines for Greek – A working* *document*. In Sonali, R., Greening, J., & Petré, L. (Authors) A comparative study of audio description guidelines prevalent in different countries (Annexe 5). London, UK: Royal National Institute of Blind People

Gerber, E. (2007). Seeing isn't believing: Blindness, race, and cultural Literacy. *The Senses and Society*, *2*(1), 27–40. Doi: 10.2752/174589207779996965

Goleman, D. (2006). *Emotional Intelligence: Why It Can Matter More Than IQ*. New York, NY: Bantam Books.

Graham, F. K. (1992). Attention: The heartbeat, the blink, and the brain. In B. A. Campbell, H. Hayne, & R. Richardson (Eds.), *Attention and Information Processing in Infants and Adults*. New York, NY: Psychology Press.

Gray, E. K., & Watson, D. (2007). Assessing positive and negative affect via self-report. In J. Coan, & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 171–183). New York, NY: Oxford University Press.

Greco, G. M. (2016). 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). London, UK: Palgrave Macmillan.

Grodal, T. K. (2006). The PECMA Flow: A general model of visual aesthetics. *Film Studies*, *8*(1), 1–11. Doi: 10.7227/FS.8.3

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion*, *9*(1), 87–108. Doi: 10.1080/02699939508408966

Hanzliček, Z., Matoušek, J., & Tihelka, D. (2008). Towards automatic audio track generation for Czech TV broadcasting: Initial experiments with subtitles-to-speech synthesis. *International Conference on Signal Processing Proceedings, ICSP*, 2721–2724. Doi: 10.1109/ICOSP.2008.4697710

Harmon-Jones, E. (2003). Clarifying the emotive functions of asymmetrical frontal cortical activity. *Psychophysiology*, *40*(6), 838–848. Doi: 10.1111/1469-8986.00121

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2013). Does negative affect always narrow and positive affect always broaden the mind? Considering the influence of motivational intensity on cognitive scope. *Current Directions in Psychological Science*, *22*(4), 301–307. Doi:

10.1177/0963721413481353

Harmon-Jones, E., Harmon-Jones, C., & Price, T. F. (2013). What is approach motivation?. *Emotion Review*, *5*(3), 291–295. Doi: 10.1177/1754073913477509

Harmon-Jones, E., Harmon-Jones, C., & Summerell, E. (2017). On the importance of both dimensional and discrete models of emotion. *Behavioral Sciences*, *7*(4), 66. Doi: 10.3390/bs7040066

Harrouet, C. (2016). *Multilingualism in Audio-Description: Audio-Subtitling*. London, UK: Roehampton University. Doi: 10.1002/ejoc.201200111

Heiss, C. (2004). Dubbing multilingual films: a new challenge? *Meta: Journal des Traducteurs*, *49*(1), 208–220. Doi: 10.7202/009035ar

Holsanova, J., Hildén, A., Salmson, M., & Tundell, V. K. (2015). *Audio Description and Audio Subtitles*. Stockholm, Sweden: Tundell och Salmson AB.

House of Representatives of the United States of America (1973). Rehabilitation Act (1973). House of Representatives 29 U.S.C. § 701, 26 September 1973, Washington, DC, United States of America.

Igareda, P., & Maiche, A. (2009). Audio description of emotions in films using eye tracking. In *Proceedings of the Symposium on Mental States, Emotions and their Embodiment in the AIBS convention 2009, December 1, Edinburgh, UK* (pp. 20–24). Retrieved from http://www.aisb.org.uk/convention/aisb09/Proceedings/MENTALSTATES/MentalStates.htm (January 2019)

Igareda, P., & Matamala, A. (2012). Variations on the Pear Tree experiment: Different variables, new results? *Perspectives: Studies in Translation Theory and Practice*, *20*(1), 103–123. Doi: 10.1080/0907676X.2011.632684

IJsselsteijn, W. a., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: Concept, determinants and measurement. In B. E. Rogowitz, & T. N. Pappas (Eds.), *Proceedings of the SPIE: Human Vision and Electronic Imaging V* (Vol. 31, pp. 520–529). Doi: 10.1117/12.387188

IMBd. (Online). Retrieved from https://www.imdb.com/ (December 2018)

Instrucció general del Consell de l'Audiovisual de Catalunya sobre l'accessibilitat als continguts audiovisuals televisius (2013). Consell de l'Audiovisual de Catalunya 19/2013, Barcelona, Spain.

International Organization for Standardization & International Electrotechnical Commission (2015). *Information technology - User Interface Component Accessibility - Part 21: Guidance on Audio Descriptions* (ISO/IEC TS 20071-21). Retrieved from https://www.iso.org/standard/63061.html (January 2019)

International Telecommunication Union (Focus Group on Audiovisual Media Accessibility) (2013). *Part 5: Final Report of Activities: Working Group B "Audio/Video Description and Spoken Captions."* Geneva, Switzerland: International Telecommunication Union

Interstate Broadcasting Treaty (2017). The Media Authorities, 1 September 2017, Berlin, Germany.

International Organization for Standardisation & International Electrotechnical Commission (2017). *Information Technology — User Interface Component Accessibility — Part 25: Guidance on the Audio Presentation of Text in Videos (Including Captions, Subtitles, and Other on Screen Text)* (ISO/IEC 20071-25:2017). Retrieved from https://www.iso.org/standard/69060.html (January 2019)

Iturregui-Gallardo, G., & Méndez-Ulrich, J. L. (In press). Towards the creation of a tactile version of the Self-Assessment Manikin (T-SAM) for the emotional assessment of visually impaired people. *International Journal of Disability, Development and Education*.

Izard, C. E. (2010). The many meanings/aspects of emotion: definitions, functions, activation, and regulation. *Emotion Review*, *2*(4), 363–370. Doi: 10.1177/1754073910374661

Izard, C. E., Dougherty, F. E., Bloxom, B. M., & Kotsch, N. E. (1974). *The Differential Emotions Scale: A Method of Measuring the Meaning of Subjective Experience of Discrete Emotions*. Nashville, TN: Vanderbilt.

Jakobson, R. (2000). On linguistic aspects of translation. In L. Venuti (Ed.), *The Translation Studies Reader* (pp. 232–239). London, UK and New York, NY: Routledge. Retrieved from http://www.stanford.edu/~eckert/PDF/jakobson.pdf (December 2018)

James, W. (1884). What is an emotion? *Mind*, *9*, 188–205.

Jankowska, A. (2015). *Translating Audio Description Scripts: Translation as a*

New Strategy of Creating Audio Description. Frankfurt am Main, Germany: Peter Lang.

Katz, E., & Foulkes, D. (1962). On the use of the mass media as "escape": clarification of a concept. *Public Opinion Quarterly*, *26*(3), 377. Doi: 10.1086/267111

Khalfa, S., Isabelle, P., Jean-Pierre, B., & Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience Letters*, *328*(2), 145–149. Doi: 10.1016/S0304-3940(02)00462-7

Kim, J., & André, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(12), 2067–2083. Doi: 10.1109/TPAMI.2008.26

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, *5*(3), 263–291. Doi: 10.1007/BF00993889

Klimmt, C., & Vorderer, P. (2003). Media psychology "is not yet there": Introducing theories on media entertainment to the presence debate. *Presence: Teleoperators and Virtual Environments*, *12*(4), 346–359. Doi: 10.1162/105474603322391596

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394–421. Doi: 10.1016/j.biopsycho.2010.03.010

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology*, *44*(5), 787. Doi: 10.1111/j.1469-8986.2007.00550.x

Krejtz, K., Krejtz, I., Duchowski, A., Szarkowska, A., & Walczak, A. (2012). Multimodal learning with audio description: an eye tracking study of children's gaze during a visual recognition task. In *Proceedings of the ACM Symposium on Applied Perception - SAP '12, August 3-4, Los Angeles*, 83–90. Doi: 10.1145/2338676.2338694

Kruger, J.-L., & Orero, P. (2010). Audio description, audio narration – a new era in AVT. *Perspectives: Studies in Translation Theory and Practice*, *18*(3), 141–142. Doi: 10.1080/0907676X.2010.487664

Kruger, J.-L., Soto-Sanfiel, M. T., Doherty, S., & Ibrahim, R. (2016). Towards a cognitive audiovisual translatology. In R. Muñoz Martín (Ed.), *Reembedding Translation Process Research* (pp. 171–194). Amsterdam, The Netherlands: John Benjamins. Doi: 10.1075/btl.128.09kru

Kuniecki, M., Barry, R., & Kaiser, J. (2003). The influence of emotionally relevant context on the evoked cardiac response triggered by an irrelevant stimulus. *Journal of Psychophysiology*, *17*(2), 61–68. Doi: 10.1027//0269-8803.17.2.61

Kuniecki, M., Pilarczyk, J., & Wichary, S. (2015). The color red attracts attention in an emotional context. An ERP study, *Frontiers in Human Neuroscience 9*(April), 1–14. Doi: 10.3389/fnhum.2015.00212

Künstler, I., Butkiewicz, U., & Więckowski, R. (2012). *Audiodeskrypcja - zasady tworzenia*. Warsaw, Poland: Fundacja Kultury bez Barier.

Kunzmann, U., & Grühn, D. (2005). Age differences in emotional reactivity: The sample case of sadness. *Psychology and Aging*, *20*(1), 47–59. Doi: 10.1037/0882-7974.20.1.47

Kunzmann, U., Kupperbusch, C. S., & Levenson, R. W. (2005). Behavioral inhibition and amplification during emotional arousal: A comparison of two age groups. *Psychology and Aging*, *20*(1), 144–158. Doi: 10.1037/0882-7974.20.1.144

Kurz, I. (2003). Psychophisiological stress during simultaneous interpreting: A comparison of experts and novices. *The Interpreters' Newsletter*, *12*, 51–67.

Labouvie-Vief, G., Lumley, M. A., Jain, E., & Heinze, H. (2003). Age and gender differences in cardiac reactivity and subjective emotion responses to emotional autobiographical memories. *Emotion*, *3*(2), 115–126. Doi: 10.1037/1528-3542.3.2.115

Lang, J. P., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPs): Affective Ratings of Picture and Instruction Manual (Technical Report A-8)*. Gainsville, FL: University of Florida

Lang, P. J. (1969). The mechanics of desensitization and the laboratory study of human fear. In C. M. Franks (Ed.), *Assessment and Status of the Behavior Therapies*. New York, NY: McGraw Hill.

Lang, P. J. (1988). What are the data of emotion?. In V. Hamilton, G. H. Bower, & N. H. Frijda (Eds.), *Cognitive Perspectives on Emotion and Motivation* (pp. 173–191). Dordrecht, The Netherlands: Springer. Doi: 10.1007/978-94-009-2792-6

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, *50*(5), 372–385. Doi: 10.1037/0003-066X.50.5.372

Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The psychophysiology of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions (3rd edition)* (pp. 180–195). New York, NY: Guilford Press.

Law on National Radio and Television (2005). Official Gazette X-469, 22 December 2005, Vilnius, Lithuania.

Law on Advertisement (2015). Official Gazette XII-1885, 25 June 2015, Vilnius, Lithuania.

Law on Public Information (2006). Official Gazette X-752, 11 July 2006, Vilnius, Lithuania.

Law on Broadcasting and Retransmission (2000). Official Gazette 308/2000, 4 October 2000, Bratislava, Slovakia.

Law on Electronic Media (2015). Official Gazette Mém. A – 236, 17 December 2015, Luxembourg, Luxembourg.

Law on the Provision of Information to the Public (2015). Official Gazette XII-2239, 23 December 2015, Vilnius, Lithuania.

Law on Urgent Measures for the Stimulation of Digital Terrestrial Television on the Liberation of Cable Television and the Encouragement of Pluralism (2005). Official Gazette BOE-A-2005-10069, 15 June 2005, Madrid, Spain.

Lawrence, C. A., & Barry, R. J. (2009). ERPs and the evoked cardiac response to auditory stimuli: Intensity and cognitive load effects. *Acta Neurobiologiae Experimentalis*, *69*(4), 552–559.

Lee, K. M. (2004). Presence, explicated. *Communication Theory*, *14*(1), 27–50.

Lee, S., Kim, G. J., Rizzo, A., & Park, H. (2004). Formation of spatial presence: by form or content? In *Proceedings of the 7th Annual International Workshop on Presence: Presence 2004, October 13-15, Valencia, Spain* (pp. 20–27). Retrieved from https://ispr.info/presence-conferences/previous-conferences/presence-2004/ (January 2019)

Lehr, C. (2014). *The Influence of Emotion on Language Performance: Study of a Neglected Determinant of Decision-Making in Professional Translators*.

Geneva, Switzerland: University of Geneva. Doi: 10.13097/archive-ouverte/unige:42306

Lei da Televisao e Serviços Audiovisuais a Pedido (2011). Official Gazette 8/2011, 11 April 2011, Lisboa, Portugal

Lerner, J. S., Dahl, R. E., Hariri, A. R., & Taylor, S. E. (2006). Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses. *Biological Psychiatry*, *61*(2), 253–260. Doi: 10.1016/j.biopsych.2006.08.016

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of presence inventory. *Presence: Teleoperators and Virtual Environments*, *10*(3), 282–297. Doi: 10.1162/105474601300343612

Levenson, R. W. (1988). Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. In H. R. Wagner (Ed.), *Social Physiopsychology and Emotion: Theory and Clinical Applications* (pp. 17–42). Milton, Australia: John Wiley & Sons.

Levenson, R. W., Ekman, P., Heider, K., & Friesen, W. V. (1992). Emotion and autonomic nervous system activity in the Minangkabau of West Sumatra. *Journal of Personality and Social Psychology*, *62*(6), 972–988. Doi: 10.1037/0022-3514.62.6.972

Linder, G. (2016). *Access Services Pan European Survey 2016*. Grand-Saconnex, Switzerland: European Broadcasting Union

Lindquist, K. A., Kober, H., & Barrett, L. F. (2015). The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences*, *35*(3), 121–143. Doi: 10.1017/S0140525X11000446

Lithuanian Journalists and Publishers Ethics Code (2005). Lithuanian Journalists and Publishers, 15 April 2005, Vilnius, Lithuania.

Ljunglöf, P., Derbring, S., & Olsson, M. (2012). A free and open-source tool that reads movie subtitles aloud. In *Proceedings of the 3rd Workshop on Speech and Language Processing for Assistive Technologies* (pp. 1–4). Retrieved from https://dl.acm.org/citation.cfm?id=2392856 (January 2018)

Loi relative à la liberté de communication (2016). Official Gazette 86-1067, 6 July 2016, Paris, France.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence.

Journal of Computer-Mediated Communication, *3*(2). Doi: 10.1111/j.1083-6101.1997.tb00072.x/full

Lombard, M., & Jones, M. T. (2015). Defining presence. In F. Biocca, M. Lombard, J. Freeman, W. A. Ijsselsteijn, & R. J. Schaevitz (Eds.), *Immersed in Media: Telepresence Theory, Measurement & Technology* (pp. 13–34). Cham: Springer International Publishing.

Luminet, O., Rimé, B., Bagby, R. M., & Taylor, G. J. (2004). A multimodal investigation of emotional responding in alexithymia. *Cognition and Emotion*, *18*(6), 741–766. Doi: 10.1080/02699930341000275

Lyons, E. J., Tate, D. F., Ward, D. S., Ribisl, K. M., Bowling, J. M, & Kalyanaraman, S. (2014). Engagement, enjoyment, and energy expenditure during active video game play. *Health Psychology*, *33*(2), 174–181. Doi: 10.1037/a0031947

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, *25*(2), 141–158.

Martínez Sanchez, F. (1996). Adaptación española de la escala de Alexitimia de Toronto (TAS-20). *Clínica y Salud*, *7*(1), 1–13.

Maszerowska, A. (2014). The visual multiplicity of films and its implications for audio description: A case study of the film *What Dreams May Come*. *Linguistica Antverpiensia. New Series – Themes in Translation Studies*, *13(1)*, 292–309.

Maszerowska, A., Matamala, A., & Orero, P. (2014). *Audio Description: New Perspectives Illustrated*. Amsterdam, The Netherlands: John Benjamins.

Matamala, A. (2014). Audio describing text on screen. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio Description: New Perspectives Illustrated* (pp. 103–120). Amsterdam, The Netherlands: John Benjamins.

Matamala, A. (2018). Voice-over: Practice, research and future prospects. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation*. Abingdon, UK: Routledge.

Matamala, A., & Orero, P. (2007). Accessible opera in Catalan: Opera for all. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 201–214).

Amsterdam, The Netherlands: Rodopi.

Matamala, A., Perego, E., & Bottiroli, S. (2017). Dubbing versus subtitling yet again? An empirical study on user comprehension. *Babel*, *63*(3), 423–441.

Matousek, J., & Vit, J. (2012). Improving automatic dubbing with subtitle timing optimisation using video cut detection. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), March 25-30, Kyoto, Japan* (pp. 2385–2388). Doi: 10.1109/ICASSP.2012.6288395

Matthews, G., Jones, D. M., & Graham Chamberlain, A. (1990). Refining the measurement of mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology*, *81*, 17–42. Doi: 10.1111/j.2044-8295.1990.tb02343.x

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209–237. Doi: 10.1080/02699930802204677

McNair, D., Lorr, M., & Droppleman, L. (1971). *Manual for the Profile of Mood States*. San Diego: Educational and Industrial Testing Service.

Media Access Australia. (2012). *Receiver-Mixed and Broadcast-Mixed Audio Description*. Sydney, Australia: Media Access Australia. Retrieved from https://mediaaccess.org.au/television/audio-description-on-tv/receiver-mixed-and-broadcast-mixed-audio-description (January 2019)

Media Accessibility Platform (n.d.). *Accessometer*. Retrieved from https://mapaccess.uab.cat/accessometer (January 2019)

Media Decision (2010). Official Gazette WJZ / 88447 (8238), 17 December 2008, The Hague, Netherlands.

Media Services Act (2014). Official Gazette RT I, 1 January 2014, Tallinn, Estonia.

Meehan, M., Insko, B., Whitton, M. C., & Brooks, F. (2002). Physiological measures of presence in virtual environments. *ACM Transactions on Graphics (TOG), 21*(3), 645–652. Doi: 10.1145/566570.566630

Memòria anual d’activitats (2016). Corporació Catalana de Mitjans Audiovisuals (Council of Ministers of the 21 July 2017). Barcelona, Spain: Corporació Catalana de Mitjans Audiovisuals

Méndez-Ulrich, J. L., Prats-Basset, L., Yagüe, F., & Sanz, A. (2016).

Percepción de control, afrontamiento y ajuste psicológico a la discapacidad visual. *Ansiedad y Estrés*, *22*(2–3), 55–61. Doi: 10.1016/j.anyes.2016.09.002

Michailidis, L., Balaguer-Ballester, E., & He, X. (2018). Flow and immersion in video games : the aftermath of a conceptual challenge. *Frontiers in Psychology*, *9*(September), 1–8. Doi: 10.3389/fpsyg.2018.01682

Miesenberger, K., Klaus, J., & Zagler, W. (2002). Computers helping people with special needs. In K. Miesenberger, J. Klaus, & W. Zagler (Eds.), *Proceedings of Computers Helping People with Special Needs, 8th International Conference on Computers for Handicapped Persons, July 15-20, Linz, Austria.* Berlin, Germany: Springer.

Mihkla, M., Hein, I., Kiissel, I., Räpp, A., Sirts, R., & Valdna, T. (2013). Subtiitrite helindamine – kas, kuidas, kellele ja milleks? *Language and Literature (Keel Ja Kirjandus)*, (11–72013), 819–827. Retrieved from https://www.europeana.eu/portal/en/record/9200419/BibliographicResource_3000125949062.html (January 2019)

Mihkla, M., Hein, I., Kiissel, I., Räpp, A., Sirts, R., & Valdna, T. (2014). A system of spoken subtitles for estonian television. *Human Language Technologies - the Baltic Perspective*, *268*, 19–26. Doi: 10.3233/978-1-61499-442-8-19

Mikropoulos, T. A., Tzimas, E., & Dimos, G. E. (2004). Objective presence measures through electric brain activity. In *Proceedings of the 7th Annual International Workshop on Presence: Presence 2004*, *October 23-15, Valencia, Spain* (pp. 259–265).

Mikul, C. (2010). *Audio Description Background Paper*. Sydney, Australia: Media Access Australia.

Milligan, B., & Fels, D. (2013). *Descriptive Video Production and Presentation. Best Practices Guide for Digital Environments*. Toronto & Ottawa, Canada: Media Access Canada.

Miron, D. (2006). Emotion and cognition in entertainment. In J. Bryant, & P. Vorderer (Eds.), *Psychology of Entertainment* (pp. 343–364). Mahwah, NJ: Erlbaum.

Moledo, A. (2017). Users' perspective on a converged audiovisual environment. In *Advanced Research Seminar in Audio Description 2017, March 16-17, Barcelona, Spain*. Retrieved from http://grupsderecerca.uab.cat/

arsad/content/programme (January 2019)

Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods*, *46*(3), 887–903. Doi: 10.3758/s13428-013-0405-3

Montoya, P., Campos, J. J., & Schandry, R. (2005). See red? Turn pale? Unveiling emotions through cardiovascular and hemodynamic changes. *The Spanish Journal of Psychologyournal of Psychology*, *8*(01), 79–85. Doi: 10.1017/S1138741600004984

Morisset, L., & Gonant, F. (2008). *Charte de l'audiodescription*. Paris, France: Conseil Supérieur de l'Audiovisuel. Retrieved from https://www.csa.fr/Arbitrer/Espace-juridique/Les-relations-du-CSA-avec-les-editeurs/Chartes/Charte-de-l-audiodescription-Decembre-2008 (January 2019)

Movietalk (2014). *Spoken Subtitles and Audio Description in Cinema. Zero Project.* Stockholm, Sweden: The Swedish National Association of Dyslexia. Retrieved from https://zeroproject.org/practice/spoken-subtitles-and-audio-description-in-cinema/ (January 2019)

Nakamura, J., & Csikszentmihalyi, M. (2002). The concept of flow. In S. J. Lopez, & C. R. Snyder (Eds.), *The Concept of Flow* (pp. 89–105). London, UK: Oxford University Press.

National Action Plan on Disability 2012-2020 (2012). Federal Government of Austria. Vienna: Federal Ministry of Labour, Social Affairs and Consumer Protection.

Netflix. (n.d.). *Netflix Audio Description Syle Guide V 1.0*. Retrieved from https://partnerhelp.netflixstudios.com/hc/en-us/articles/215510667-Audio-Description-Style-Guide-v1-0 (December 2018)

New York State Web Accessibility Policy (2010). Office of Information Technology Services NYS-P08-005, New York, NY, United States of America.

Nielsen, S., & Bothe, H. H. (2008). Subpal: A device for reading aloud subtitles from television and cinema. In M. Hersh, & J. Ohene-Djan (Eds.) *Proceedings of the Conference and Workshop on Assistive Technologies for People with Vision and Hearing Impairments: Assistive Technology for*

All Ages (Vol. 415), August 28-31, Granada, Spain.

Nowlis, V. (1965). Research with the Mood Adjective Check List. In S. S. Tompkins, & C. E. Izard (Eds.), *Affect, Cognition, and Personality: Empirical Studies* (p. 464). Oxford, UK: Springer.

O'Hagan, M. (2016). Game localisation as emotion engineering: Methodological exploration. In Q. Zhang, & M. O'Hagan (Eds.), *Conflict and Communication: A Changing Asia in a Globalising World* (pp. 123–144). New York, NY: Nova.

O'Hara, J. (n.d.). *Audio Description Guidelines (for Film, Video and Television)*. London, UK: Independent Television Facilities Centre Ltd.

Office of Communications (2015). *Code on Television Access Services*. Retrieved from http://stakeholders.ofcom.org.uk/binaries/broadcast/other-codes/ctas.pdf (September 2018)

On-demand Audiovisual Media Services Act (2010). Official Gazette 132/2010, 13 April 2010, Praga, Czech Republic.

Orero, P. (2007a). Audiosubtitling: A possible solution for opera accessibility in Catalonia. *Tradterm*, *13*, 135-149. Doi: 10.11606/issn.2317-9511.tradterm.2007.47470

Orero, P. (2007b). Pioneering audio description: An interview with Jorge Arandes. *Journal of Specialised Translation*, (7), 179–189.

Orero, P. (2007c). Sampling audio description in Europe. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 111–125). Amsterdam, The Netherlands and New York, NY: Rodopi.

Orero, P. (2011). The audiodescription of spoken, tactile and written language in *Be with Me*. In A. Matamala, J.-M. Lavaur, & A. Serban (Eds.), *Audiovisual Translation in Close-up* (pp. 239–255). Bern, Switzerland: Peter Lang.

Orero, P., & Matamala, A. (2007). Accessible opera: overcoming linguistic and sensorial barriers. *Perspectives: Studies in Translation Theory and Practice*, *15*(4), 262–277.

Orero, P., & Vilaró, A. (2012). Eye tracking analysis of minor details. *Monographs in Translation and Interpreting*, *4*, 295–319.

Orero, P., & Wharton, S. (2007). The audio description of a Spanish phenomenon: *Torrente 3. The Journal of Specialised Translation*, 7, 164–178.

Orero, P., Doherty, S., Kruger, J.-L., Matamala, A., Pedersen, J., Perego, E., Romero-Fresco, P., Rovira-Esteva, S., Soler-Vilageliu, O., & Szarkowska, A. (2018). Conducting experimental research in audiovisual translation (AVT): A position paper. *The Journal of Specialised Translation*, *30*, 105–126.

Perego, E., Del Missier, F., & Bottiroli, S. (2014). Dubbing versus subtitling in young and older adults: Cognitive and evaluative aspects. *Perspectives: Studies of Translation Theory and Practice*, *23*(1), 1–21. Doi: 10.1080/0907676X.2014.912343

Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The cognitive effectiveness of subtitle processing. *Media Psychology*, *13*(3), 243–272. Doi: 10.1080/15213269.2010.502873

Pereira, A. (2005). El subtitulado para sordos: estado de la cuestión en España. *Quaderns. Revista de Traducció*, *12*, 161–172.

Petrie Thomas, J. H., Whitfield, M. F., Oberlander, T. F., Synnes, A. R., & Grunau, R. E. (2012). Focused attention, heart rate deceleration, and cognitive development in preterm and full-term infants. *Developmental Psychobiology*, *54*(4), 383–400. Doi: 10.1002/dev.20597

Pfanstiehl, M. R. (2016). *The Metropolitan Washington Ear, Inc*. Retrieved from http://www.washear.org/ (January 2019)

Piferi, R. L., Kline, K. A., Younger, J., & Lawler, K. A. (2000). An alternative approach for achieving cardiovascular baseline: Viewing an aquatic video. *International Journal of Psychophysiology*, *37*(2), 207–217. Doi: 10.1016/S0167-8760(00)00102-1

Pishghadam, R., Firooziyan, A., & Esfahani, P. (2017). Introducing emotioncy as an effective tool for the acceptance of Persian neologisms. *Language Related Research*, *8(5)*, 79-105.

Plantinga, C. (2009). Emotion and Affect. In P. Livingstone, & C. Plantinga (Eds.), *The Routledge Companion to Philosophy and Film* (pp. 86–96). London, UK and New York, NY: Routledge. Doi: 10.4324/9780203879320

Plutchik. (2009). Emotions: A general psychoevolutionary theory. In K. R. Scherer, & P. Ekman (Eds.), *Approaches to Emotion*. New York, NY and

London, UK: Psychology Press.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of Emotion (Vol. 21)*. 3–33. Doi: 10.1016/B978-0-12-558701-3.50007-7

Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots. *American Scientist, 89(4),* 344-350. Doi: 10.1511/2001.4.344

Polyglot Nerd (n.d.). Retrieved from http://polyglotnerd.tumblr.com/ (December 2018)

Presidential Decree on the harmonisation with Directive 2010/13/EU (2010). Official Gazette 109/2010, 4 November 2010, Athens, Greece.

Proposal for a directive of the european parliament and of the council on the approximation of the laws, regulations and administrative provisions of the Member States as regards the accessibility requirements for products and services (2018). European Commision 15174/18 ADD 1 2015/0278, 12 December 2018, Brussels, Belgium.

Public service contract between DR and the Danish Minister for Culture 2019-2023 (2018). Danish Ministry of Culture, 18 September 2018, Copenhagen, Denmark.

Puigdomènech, L., Matamala, A., & Orero, P. (2008). The making of a protocol for opera audio description. In L. Pegenaute, J. de Cesaris, M. Tricás, & E. Bernal (Eds.), *Actas del III Congreso Internacional de la Asociación Ibérica de Estudios de Traducción e Interpretación. La traducción del futuro: mediación lingüístca y cultural en el siglo XXI. Barcelona, 22-24 March 2007* (pp. 381–392). Barcelona: PPU

Pujol, J., & Orero, P. (2007). Audio description precursors: Ekphrasis, film narrators and radio journalists. *Translation Watch Quarterly*, *3*(2), 49–60.

Radio and Television Broadcasters Law (2016). Official Gazette 7(I)98, 27 October 2016, Athens, Greece.

Radio and Television of Slovenia Act (2005). Official Gazette 96/2005, 28 October 2005, Ljubljana, Slovenia.

Rai, S., Greening, J., & Leen, P. (2010). *A Comparative Study of Audio Description Guidelines Prevalent in Different Countries*. London, UK: Royal Institute of the Blind People. Retrieved from http://www.rnib.org.uk/professionals/Documents/International_AD_Standar ds_comparative study_2010.doc (December 2018)

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, *61*(1), 5–18. Doi: 10.1016/j.ijpsycho.2005.10.024

Ramos Caro, M. (2016). *La traducción de los sentidos*. Munich: Lincom GmbH.

Ramos, M. (2015). The emotional experience of films: Does audio description make a difference?. *The Translator*, *21*(1), 68–94. Doi: 10.1080/13556509.2014.994853

Ramos, M., & Rojo, A. M. (2014). "Feeling" audio description. Exploring the impact of AD on emotional response. *Translation Spaces*, *3*, 133–150. Doi: 10.1075/ts.3.06ram

Ratan, R. (2013). Self-Presence, Explicated: Body, Emotion, and Identity Extension into the Virtual Self. In R. Luppicini (Ed.), *Handbook of Research on Technoself: Identity in a Technological Society* (pp. 322–336). Hershey, PA: Information Science Reference. Doi: 10.4018/978-1-4666-2211-1.ch018

Reeves, B., Lang, A., Kim, E. Y., & Tatar, D. (2009). The effects of screen size and message content on attention and arousal. *Media Psychology*, *1*(1), 49–67. Doi: 10.1207/s1532785xmep0101_4

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Stanford, UK: CSLI Publications.

Regenbrecht, H. T., Schubert, T. W., & Friedmann, F. (1998). Measuring the sense of presence and its relations to fear of heights in virtual environments. *International Journal of Human-Computer Interaction*, *10*(3), 233–249. Doi: 10.1207/s15327590ijhc1003_2

Remael, A. (2005). *Audio Description for Recorded TV, Cinema and DVD Experimental Style Sheet* (Unpublished document)

Remael, A. (2012). Audio description with audio subtitling for Dutch multilingual films: Manipulating textual cohesion on different levels. *Meta: Journal Des Traducteurs*, *57*(2), 385–407. Doi: 10.7202/1013952ar

Remael, A. (2014). Combining audio description with audio subtitling. In A. Remael, N. Reviers, & G. Vercauteren (Eds.), *ADLAB Audio Description*

Guidelines. Retrieved from http://www.adlabproject.eu/Docs/adlab book/index.html#combining-ad (December 2018)

Remael, A., Reviers, N., & Vercauteren, G. (2014). *ADLAB Audio Description Guidelines.* Retrieved from http://www.adlabproject.eu/Docs/adlab%20book/index.html (December 2018)

Reviers, N., & Remael, A. (2015). Recreating multimodal cohesion in audio description: A case study of audio subtitling in Dutch multilingual films. *New Voices in Translation Studies*, *13*(1), 50–78.

Ritz, T., Steptoe, A., DeWilde, S., & Costa, M. (2000). Emotions and stress increase respiratory resistance in asthma. *Psychosomatic Medicine*, *62*(3), 401–12.

Riva, G. (2011). Presence, actions and emotions: A theoretical framework. *Annual Review of CyberTherapy and Telemedicine*, *9*(1), 2–5.

Rojo, A., Ramos, M., & Valenzuela, J. (2014). The emotional impact of translation: A heart rate study. *Journal of Pragmatics*, *71*(September), 31–44. Doi: 10.1016/j.pragma.2014.07.006

Romero-Fresco, P., & Fryer, L. (2014a). Audiointroductions. In A. Matamala, A. Maszerowska, & P. Orero (Eds.), *Audio Description: New Perspectives Illustrated* (p. 216). Amsterdam, The Netherlands: John Benjamins.

Romero-Fresco, P., & Fryer, L. (2014b). Could audio described films benefit from audio introductions? An audience response study. *Journal of Visual Impairment & Blindness*, *107*(4), 287–295.

Rottenberg, J. (2003). When emotion goes wrong: Realizing the promise of affective science. *Clinical Psychology: Science and Practice*, *10*(2), 227–232. Doi: 10.1093/clipsy/bpg012

Rottenberg, J., Ray, R. D., & Gross, J. J. (2007). Emotion elicitation using films. In J. A. Coan, & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 9–28). New York, NY: Oxford University Press.

Round Table (Online). *Round Table on Information Access for People with Print Disabilities*. Retrieved from http://printdisability.org/ (December 2018)

Rovira-Esteva, S., & Tor-Carroggio, I. (2018). Serveis d'accessibilitat sensorial a les televisions que emeten en català: situació actual i propostes de futur. *Quaderns del CAC*, *44*(XXI), 71–80.

Rowe, T. L. (1960). The English dubbing text. *Babel*, *6*(3), 116–120.

Rozin, P. (1999). Preadaptation and the puzzles and properties of pleasure. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-Being: The Foundations of Hedonic Psychology* (pp. 109–133). New York, NY: Russell Sage.

Rozin, P., Guillot, L., Fincher, K., Rozin, A., & Tsukayama, E. (2013). Glad to be sad, and other examples of benight masochism. *Judgment and Decision Making*, *8*(4), 439–447.

Ruggiero, T. E. (2000). Uses and gratifications theory in the 21st century. *Mass Communication & Society*, *3*(1), 3–37. Doi: 10.1207/S15327825MCS0301_02

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. Doi: 10.1037/h0077714

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, *57*(3), 493–502. Doi: 10.1037/0022-3514.57.3.493

Sánchez Mompeán, S. (2016). "It's not what they said; it's how they said it": A corpus-based study on the translation of intonation for dubbing. In A. M. Rojo López, & N. Campos Plaza (Eds.), *Interdisciplinarity in Translation Studies. Theoretical Models, Creative Approaches and Applied Methods*. Bern, Switzerland and New York, NY: Peter Lang. Doi: 10.3726/978-3-0351-0954-2

Schaefer, A., Nils, F., Philippot, P., & Sanchez, X. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion*, *24*(7), 1153–1172. Doi: 10.1080/02699930903274322

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. Doi: 10.1177/0539018405058216

Schlosberg, H. (1954). The psychological review. *The Psychological Review*, *63*(2), 81–88.

Schramm, H., & Wirth, W. (2010). Exploring the paradox of sad-film enjoyment: The role of multiple appraisals and meta-appraisals. *Poetics*, *38*(3), 319–335. Doi: 10.1016/j.poetic.2010.03.002

Schreiber, P. A. (1987). Prosody and structure in children's syntactic processing. In R. Horowitz, & S. J. Samuels (Eds.), *Comprehending Oral and Written Language* (pp. 243–270). San Diego, CA: Academic Press.

Schwanenflugel, P. J., Hamilton, A. M., Wisenbaker, J. M., Ruhn, M. R., & Stahl, S. A. (2009). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology*, *96*(1), 119–129. Doi: 10.1037/0022-0663.96.1.119.Becoming

Sepielak, K. (2016a). Synchronization techniques in multilingual fiction: Voiced-over films in Poland. *International Journal of Communication*, *10*, 1054–1073.

Sepielak, K. (2016b). The effect of subtitling and voice-over on content comprehension and language identification in multilingual movies. *International Journal of Sciences: Basic and Applied Research*, *25*(1), 166–182.

Serrano-Puche, J. (2016). Internet and emotions: New trends in an emerging field of research / Internet y emociones : nuevas tendencias en un campo de investigación emergente. *Comunicar*, *24*(46), 19–26. Doi: 10.3916/C46-2016-02

Service Provision Agreement Rai 2018-2022 (2018). Official Gazette, 28 April 2017, Ministerio dello Sviluppo, Rome, Italy.

Sherry, J. J. L. (2004). Flow and media enjoyment. *Communication Theory*, *14*(4), 328–347. Doi: 10.1111/j.1468-2885.2004.tb00318.x

Silverman, C., & Fels, D. I. (2002). Emotive captioning in a digital world. In K. Miesenberger, J. Klaus, & W. Zagler (Eds.), *Proceedings of Computers* 292–294.

Simons, R., Detenber, B., & Roedema, T. (1999). Emotion processing in three systems: The medium and the message. *Psychophysiology*, *36*, 619–627.

Smith, C. A., & Lazarus, R. S. (1993). Appraisal components, core relational themes, and the emotions. *Cognition & Emotion*, *7*(3–4), 233–269. Doi: 10.1080/02699939308409189

Snyder, J. (2010). *Audio Description Guidelines and Best Practices*. Snyder, J. (Ed.). Alexandria, VA: The American Council of the Blind

Sohn, J. H., Sokhadze, E., & Watanuki, S. (2001). Electrodermal and cardiovascular manifestations of emotions in children. *Journal of Physiological Anthropology and Applied Human Science*, *20*(2), 55–64.

Doi: 10.2114/jpa.20.55

Sokolov, E. N., & Boucsein, W. (2000). A psychophysiological model of emotion space. *Integrative Physiological and Behavioral Science*, *35*(2), 81–119. Doi: 10.1007/BF02688770

Spielberg, C. D., Gorusch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*. Doi: 10.3758/s13428-017-0962-y

Stallard, G. (2003). *Standardisation Requirements for Access to Digital TV and Interactive Services by Disabled People. (Final Report to CENELEC on TV for All)*. Chandler's Ford, UK: European Committee for European Standardization. Retrieved from http://www.acessibilidade.net/tdt/FinalreportTVforAll.pdf (January 2019)

Stearns, P. N., & Stearns, C. Z. (1985). Emotionology: Clarifying the history of emotions and emotional standards. *The American Historical Review*, *90*(4), 813–836.

Stemmler, G. (2004). Physiological processes during emotion. In P. Philippot, & R. S. Feldman (Eds.), *The Regulation of Emotion* (pp. 33–70). Mahwah, NJ and London, UK: Lawrence Erlbaum Associates. Doi: 10.4324/9781410610898

Sternberg, M. (1981). Polylingualism as reality and translation as mimesis. *Poetics Today*, *2*(4), 221–239.

Szarkowska, A. (2011). Text-to-speech audio description: Towards wider availability of AD. *The Journal of Specialised Translation*, *15*, 142–162.

Szarkowska, A., & Jankowska, A. (2012). Text-to-speech audio description of voiced-over films. A case study of audio described *Volver* in Polish. In E. Perego (Ed.), *Emerging Topics in Translation: Audio Description* (pp. 81–98). Trieste, Italy: EUT Edizioni Università di Trieste.

Szarkowska, A., & Mączyńska, M. (2011). Text-to-speech audio description with audio subtitling to a non-fiction film. a case study based on *La Soufriere* by Werner Herzog. Paper presented at the Audiovisual Translation Research

Lab, London, UK, 1 July 2011.

Szarkowska, A., & Walczak, 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* (pp. 209–233). Bern, Switzerland: Peter Lang.

Szarkowska, A., Zbikowska, J., & Krejtz, I. (2013). Subtitling for the deaf and hard of hearing in multilingual films. *International Journal of Multilingualism*, *10*(3), 292–312.

Szymańska, B., & Strzymiński, T. (2010). *Standardy tworzenia audiodeskrypcji do produkcji do audiowizualnych*. Bialystok, Poland: Fundacja Audiodeskrypcja.

Tactile Book Advancement Group (2006). *Communicating Tactile Story Ideas to Young Visually Impaired Children*. Retrieved from http://www.tactilebooks.org/making/telling-touch.pdf (December 2018)

Tanton, N. (2015). *ATVoice-overD Working Group on Access Services (WGAS). 2015 Report*. London, UK: Office of Communications. Retrieved from https://www.ofcom.org.uk/__data/assets/pdf_file/0025/83167/working_group_on_access_services_first_report_july_2014.pdf (January 2019)

Taylor, G. J., Ryan, D. P., & Bagby, R. M. (1985). Toward the development of a new self-report alexithymia scale. *Psychotherapy and Psychosomatics*, *44*(4), 191–9.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*(1), 67–85. Doi: 10.1037/0033-2909.110.1.67

Thayer, R. E. (1967). Measurement of activation through self-report. *Psychological Reports*, *20*(2), 663–678.

Thayer, R. E. (1986). Activation-Deactivation Adjective Check List: Current overview and structural analysis. *Psychological Reports*, *58*(2), 607–614.

The Audiovisual Law (2001). Official Gazette 504/2002, 11 July 2002, Bucarest, Romania.

The Braille Authority of North America. (2011). *Guidelines and Standards for Tactile Graphics*. Baltimore, MD: The Braille Authority of North America Retrieved from http://www.brailleauthority.org/tg/web-manual/index.html

(February 2018)

Thrane, L. K. (2013). *Text-to-Speech on Digital TV: An Explanatory Study of Spoken Subtitles on DR1Syn*. Copenhagen, Denmark: ITU University of Copenhagen. Retrieved from http://www.respeaking.net/Barcelona%202013/Lisbet%20Kvistholm%20Thrane%20-%20Text-tospeech%20on%20Digital%20TV%20-%20A%20Case%20Study%20of%20Spoken%20Subtitles%20on%20DR1Syn.pdf (January 2019)

Ticineto Clough, P., & Halley, J. (2007). *The affective turn: theorizing the social*. Durham, UK: Duke University Press.

TraFilm (2017). *TraFilm Project Database*. Retrieved from http://gallery.trafilm.net/ (December 2018)

TransMedia Catalonia (2019). *NEA Results by Gonzalo Iturregui-Gallardo: The Tactile SAM.* Retrieved from https://www.youtube.com/watch?v=mcA_s4sVpaA (January 2019)

Tremayne, P., & Barry, R. J. (2001). Elite pistol shooters: physiological patterning of best vs. worst shots. *International Journal of Psychophysiology*, *41*(1), 19–29. Doi: 10.1016/S0167-8760(00)00175-6

Tsai, J. L., Levenson, R. W., & Carstensen, L. L. (2000). Autonomic, subjective, and expressive responses to emotional films in older and younger Chinese Americans and European Americans. *Psychology and Aging*, *15*(4), 684–693. Doi: 10.1037/0882-7974.15.4.684

Tsai, J. L., Pole, N., Levenson, R. W., & Muñoz, R. F. (2003). The effects of depression on the emotional responses of Spanish-speaking Latinas. *Cultural Diversity and Ethnic Minority Psychology*, *9*(1), 49–63. Doi: 10.1037/1099-9809.9.1.49

Tsaousi, A. (2015). Making sound accessible: The labelling of soundeffects in subtitling for the deaf and hard-ofhearing. *Hermeneus*, *17*, 233–252.

Twenty-First Century Communications and Video Accessibility Act (2010). House of Representative of the United States of America 111-260, 8 October 2010, Washington, DC, United States of America

Udo, J. P., Acevedo, B., & Fels, D. I. (2010). Horatio audio-describes Shakespeare's Hamlet: Blind and low-vision theatre-goers evaluate an unconventional audio description strategy. *British Journal of Visual Impairment*, *28*(2), 139–156. Doi: 10.1177/0264619609359753

Udo, J. P., & Fels, D. (2009). Re-fashioning fashion: an exploratory study of a live audio described fashion show. *Universal Access in the Information Society*, *9*(1), 63–75. Doi: 10.1007/s10209-009-0156-1

Uhrig, M. K., Trautmann, N., Baumgärtner, U., Treede, R.-D., Henrich, F., Hiller, W., & Marschall, S. (2016). Emotion elicitation: A comparison of pictures and films. *Frontiers in Psychology*, *7*(February). Doi: 10.3389/fpsyg.2016.00180

UK Association for Accessible Formats (2012). *Creating clear print and large print documents*. Cwmbran, UK: UK Association for Accessible Formats

Usoh, M., Arthur, K., Whitton, M., Bastos, R., Steed, A., Brooks, F., & Slater, M. (1999). The Visual Cliff revisited: A virtual presence study on locomotion. *Proceedings of the Second International Workshop on Presence*, *April 6-7, Colchester, UK.*

Verboom, M., Crombie, D., Dijk, E., & Theunisz, M. (2002). Spoken subtitles: Making subtitled TV programmes accessible. In K. Miesenberger, J. Klaus, & W. Zagler (Eds.), *Computers Helping People with Special Needs. 8th International Conference, ICCHP 2002, July 15-20, Linz, Austria* (pp. 295–302).

Vercauteren, G. (2012). A narratological approach to content selection in audio description: Towards a strategy for the description of narratological time. *Monographs in Translation and Interpreting*, (4), 207–231. Doi: 10.6035/MonTI.2012.4.9

Vianna, E. P. M., & Tranel, D. (2006). Gastric myoelectrical activity as an index of emotional arousal. *International Journal of Psychophysiology*, *61*(1), 70–76. Doi: 10.1016/j.ijpsycho.2005.10.019

Video Description: Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010 (2017). Federal Communications Commision 2017-15526, 8 October 2017, Washington, DC, United States of America.

Vilaró, A., Duchowski, A., Orero, P., Grindinger, T., Tetreault, S., & Di Giovanni, E. (2012). How sound is the Pear Tree Story? Testing the effect of varying audio stimuli on visual attention distribution. *Perspectives: Studies in Translation Theory and Practice*, *20*(1), 55–65.

Vision Australia (n.d.). *What is Print Disability?* Retrieved from https://www.visionaustralia.org/services/print-accessibility/what-is-print-disability (January 2019)

Voellmer, E., & Zabalbeascoa, P. (2014). How multilingual can a dubbed film be? Language combinations and national traditions as determining factors. *Linguistica Antverpiensia, New Series*, *13*, 232–250.

Vorderer, P., & Hartmann, T. (2008). Entertainment and enjoyment as media effects. In J. Bryant, & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (pp. 532–550). London, UK: Routledge.

Walczak, A. (2010). *Audio Description for Children. A Case Study of Text-to-Speech Audio Description of Educational Animation Series* Once Upon a Time... Life. Warsaw, Poland: University of Warsaw. (Unpublished MA Thesis)

Walczak, A. (2017). *Immersion in Audio Description: The Impact of Style and Vocal Delivery on Users' Experience.* Bellaterra, Spain: Universitat Autònoma de Barcelona. Retrieved from https://www.tdx.cat/handle/10803/402401 (January 2019)

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

Walczak, A., & Fryer, L. (2018). Vocal delivery of audio description by genre: measuring users' presence. *Perspectives: Studies in Translation Theory and Practice*, 1–15. Doi: 10.1080/0907676X.2017.1298634

Watson, D. (2000). *Mood and Temperament*. New York, NY: Guilford Press

Watson, D., & Clark, L. A. (1999). *The PANAS-X: Manual for the Positive and Negative Affect Schedule - Expanded Form*. Iowa City, IA: Department of Psychological & Brain Sciences Publications.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. Doi: 10.1037/0022-3514.54.6.1063

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235. Doi: 10.1037/0033-2909.98.2.219

Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., & Groner, R. (2008). Playing online games against computer- vs. human-controlled opponents:

Effects on presence, flow, and enjoyment. *Computers in Human Behavior*, *24*(5), 2274–2291. Doi: 10.1016/j.chb.2007.11.002

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, *26*(4), 557–580. Doi: 10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

Widen, S. C., & Russell, J. A. (2010). Descriptive and prescriptive definitions of emotion. *Emotion Review*, *2*(4), 377–378. Doi: 10.1177/1754073910374667

Wikipedia (n.d.). *Multilingual Films*. Retrieved from https://en.wikipedia.org/wiki/Category:Multilingual_films (January 2019)

Wilken, N., & Kruger, J.-L. (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. Doi: 10.1556/084.2016.17.2.6

Wilson, G. M., & Sasse, M. A. (2006). Do users always know what's good for them? Utilising physiological responses to assess media quality. In S. McDonald, Y. Waern, & G. Cocktrn (Eds.), *People and Computers XIV — Usability or Else! Poceedings of HCI 2000* (pp. 327–339). London, UK: Springer London.

Wundt, W. (1896). *Grundriss der Psychologie [Outlines of Psychology]*. Leipzig: Wilhelm Engelmann.

Yasuda, T., Lubin, B., Kim, J., & Van Whitlock, R. (2003). The Japanese version of the Multiple Affect Adjective Checklist — Revised: Development and validation. *Journal of Clinical Psychology*, *59*(1), 93–109. Doi: 10.1002/jclp.10120

Zillman, D. (1999). Exemplification theory: Judging the whole by some of its parts. *Media Psychology*, *1*(1), 46–67. Doi: 10.1207/s1532785xmep0101

Zillmann, D., & Bryant, J. (1994). Entertainment as media effect. In J. Bryant, & D. Zillmann (Eds.), *Media Effects: Advances in Theory and Research* (pp. 437–461). Hillsdale, NJ: Erlbaum.

Zuckerman, M. (2002). Zuckerman-Kuhlman Personality Questionnaire (ZKPQ): An alternative five-factorial model. In B. de Raad, & M. Perugini (Eds.), *Big Five Assessment* (pp. 376–392). Ashland, OH: Hogrefe & Huber Publishers.

Zuckerman, M., Lubin, B., & Rinck, C. (1983). Construction of new scales for the Multiple Affect Adjective Check List. *Journal of Psychopathology and Behavioral Assessment*, *5*, 119–129.

Filmography

Affleck, B. (2012). *Argo*.

Allen, W. (2011). *Midnight in Paris*.

Allen, W. (2012). *To Rome with Love*.

Almodóvar, P. (2002). *Hable con ella*.

Almodóvar, P. (2006). *Volver*.

Ayer, D. (2014). *Fury*.

Bhansali, S. L. (2007). *Saawariya*.

Boyle, D., & Tandan, L. (2008). *Slumdog Millionaire*.

Caghan, S. (2005). *Syriana*.

Campion, J. (1993). *The piano*.

Chadha, G. (2004). *Bride & Prejudice*.

Charles, L. (2006). *Borat: Cultural Learnings of America for Make Benefit Glorious Nation of Kazakhstan*.

Dunquart, P. (2007). *Am Limit*.

Eastwood, C. (2008). *Gran Torino*.

Gavron, S. (2007). *Brick Lane*.

Gondry, M. (2006). *The Science of Sleep*.

González Iñárritu, A. (2006). *Babel*.

Johar, K. (2010). *My Name is Khan*.

Koolhoven, M. (2008). *Oorlogswinter*.

Loach, K. (2007). *It's a Free World*...

Marshall, R. (2005). *Memoirs of a Geisha*.

Murphy, R. (2010). *Eat Pray Love*.

Noyce, P. (2010). *Salt*.

Schreib, L. (2005). *Everything is illuminated*.

Scott, R. (2006). *A Good Year*.

Singer, B. (2008). *Valkyrie*.

Tarantino, Q. (2009). *Inglorious Basterds*.

van den Berg, R. (2010). *Tirza*.

van den Berg, R. (2012). *Süskind*.

Annexes

Annexes

1. Summary of Regulations, Standards and Guidelines

Country	Regulation	Standards, Guidelines and Recommendations
Australia	**Broadcasting Services Act (2018)** • For the purposes of this Act, if a television program is captioned for the deaf and hearing impaired, the captioning is taken to be part of the program.	**Audio Description Background Paper** (Mikul, 2010) • "For subtitled films, note the appearance of subtitles in the first instance, then prefix with "subtitle" or "he says", "she says", etc"
Austria	**National Action Plan on Disability 2012-2020. Strategy of the Austrian Federal Government for the Implementation of the UN Disability Rights Convention (2012)** • "The Austrian Broadcasting Corporation (ORF) has to ensure concerns of people with disability. ORF and other audiovisual media services are obliged to increase the proportion of programmes which are accessible for people with hearing disabilities and visual disabilities, particularly via sign language, subtitles, audio descriptions and easily understandable menu navigation." (There are no numbers)	
Belgium	Flemish community **Decree Amending Various Provisions of the Decree of 27 March 2009 Concerning Radio Broadcasting and Television (2018)** • 100% of news subtitled. French community **Consolidated Decree about Audiovisual Media Services (2018)** • 25% of fiction and documentary programs in prime time (13h – 24h) should be audio described in public broadcasters, and in 20% for private broadcasters in programs with more than 2.5% of audience. For public broadcasters, 15% of programs for those emissions with less than 2.5% of audience.	**Audio description for recorded TV, Cinema and DVD. Experimental stylesheet for teaching purposes** (Remael, 2005) What is to be described? • On-screen action (e.g. battle scenes), sounds or effects that cannot easily be identified, subtitles or other text on screen (e.g. graffiti), opening titles and/or end credits.
Bulgaria	**Electronic Communications Act (2007)** • "Access to conditional access systems for digital television services; access to virtual network	

	services." • "to ensure access to the program for persons with hearing and visual impairments including specialized subtitle language signs, audio description and audio subtitles"	
Canada	**Broadcasting Act (2018)** • Programming accessible by disabled persons should be provided within the Canadian broadcasting system as resources become available for the purpose.	**Descriptive Video Production and Presentation of Best Practices Guide for Digital Environments** (Miligan & Fels, 2013) Subtitles • Subtitles are sometimes used to provide translation when characters speak in a different language from the primary language of t show. In these instances, the describer should read the translation after stating that a subtitle appears. The recommended practice to add a description with the sound effects at the beginning of the video if there are subtitles that the describer will be reading later in the video. This is considered to be "describing the screen real estate". **Described Video Best Practices. Artistic and Technical Guideline.** Accessible Media Inc and The Canadian Association of Broadcasters (2014) • Recommendations: Captions and subtitles are used to visually display information on the screen that may not be available in the original audio track, such as the translation of a foreign language, the passage of time or otherwise quiet dialogue. Other requirements may exist for the proper pronunciation of a foreign language. The inclusion of the description of these elements should be evaluated based upon the composition of the original audio track, the relevancy of the content and whether it develops the material as a whole. • Techniques: When confronted with text on the screen, say the associated word, such as "title," "subtitle," "caption," "logo," "credits," "end credits," "sign reads" or "sign on the wall," followed by the text being displayed or a description of its content. If the text appears in an abnormal fashion, such as through short-hand used in a text message, describe the context of the abnormal text if time permits.

Catalonia	**General instruction of the Catalan Audiovisual Council on access to TV audiovisual contents (2013)** • Definition of AD: "access service to communication based on the series of techniques and abilities applied with the aim of balancing the lack of perception of the visual part found in any kind of message, by providing with the suitable aural information that translates or explains, in a way that the person with visual disability could perceive this message as a harmonic unity and in the most similar way to that perceived by a sighted person." • 10 hours per week.	**Basis for a Future Audio Description Protocol in the Catalan Ambit** Domènech, Orero and Matamala (2007) • What has to be described? - Subtitles and surtitles • 6.6. Audio description of written elements b) Subtitles. As a matter of fact, the apparition of subtitles in an audiovisual content is often derived from a sporadic or continued presence of a linguistic code different from the principal linguistic code that is presumably shared by the target audience. That way, in section 6.7the approach to other languages in a work concerning audio description is commented. As long as audiovisual products are concerned, Ofcom (2006) and O'Hara (2004) offer different alternative, without making explicit, though, which is preferable. On the one hand, the audio descriptor can announce it is a subtitle; for example: "Pere says in polish:" + reading aloud of the text subtitled in Catalan, or "A subtitle reads:" + reading of the subtitle. On the other hand, a different
Croatia	**Electronic Media Act (2013)** • It will encourage providers of AV media services to make services accessible. **Act on Croatian Radio-television (2012)** • Croatian radio-television requires to gradually make accessible their AV media services.	
Cyprus	**Persons with Disabilities Act (2000)** • News bulletin is comprehensible by the deaf. **Radio and Television Broadcasters Law (2016)** • Specific ways to achieve gradual increase in the percentage of its shows (one-year timeline) that will be accessible to persons with visual or hearing disability by at least 5%.	
Czech Republic	**Act on Czech Television (1992)** • Says that at least a 2% of the programmes must be accessible to bling people.	

	On-demand Audiovisual Media Services Act (2010) Section 6 • (4) An on-demand audiovisual media service provider shall, where appropriate, provide a programme with open or closed captioning) or with interpreting into Czech sign language) for persons with hearing impairments, and with a sound track for persons with visual impairments, if available, or shall otherwise ensure that certain programmes provided via an on-demand audiovisual media service are accessible to persons with hearing impairments and persons with visual impairments.	
Denmark	**Contract between DR and Danish Minister for Culture,2019-2023 (2018)** • "removing barriers for people who find it difficult to hear, see, read or understand, so that they have better opportunities to benefit from DR's public service content" • "DR must increase the knowledge of DR's accessibility services in the relevant target groups" • "DR must be in regular dialogue with relevant disability organizations about the accessibility work"	
Estonia	**Media Services Act (2014)** • "23. Access to people with visual or hearing disability. Audiovisual media service provider shall make the service accessible to people with a visual or hearing disability using for that purpose, among other, the supplying of the program with subtitles, sign language translation, separate audio channels, teletext and other ancillary services that enable people with a visual or hearing disability to use the provided service. "	
EU	**Directive on Coordination of Audiovisual Services among the Member States (Audiovisual Media Services Directive (2010)** • (46) The right of persons with a disability and of the elderly to participate and be integrated in the social and cultural life of the Union is inextricably linked to the provision of accessible audiovisual media services. The means to achieve accessibility should include, but need not be limited to, sign language, subtitling, audio-description and easily understandable menu navigation."	**ADLAB Project** Remael (2014) • Explanation of what AST are • Possible scenarios for AST • She says that the way AST is provided with audio-described films is not regulated and will vary from country to country • She gives recommendations on how to include AST in the AD script, which is sometimes already done and impossible to change: - Indicate the speaker - Such subtitles are often recorded as voice-over so the audience will probably still have access to the

		sound of the original dialogues for character identification. If there is no time to indicate who is speaking this way not be a problem in that case. - If expanded subtitling is required, made-to-measure for the AST, translate the ST dialogues as dialogue without applying the condensation that is usual for subtitling and write your AD accordingly. If the expanded subtitling is going to be recorded dubbing style, make sure that your AD identifies who is speaking. The fewer voices are used for the recording, the more attention may need to be paid to character identification in the AD. - If you are the subtitler but your film has already been audio described: make sure that the AST contains all the info required to interact with the existing AD and that together, they cover the info that the sighted audience would get form combining visuals and dialogue.
Finland	**Act on Television and Radio Operations (2010)** • Section 19. Making programmes accessible to people with a visual or hearing impairment. • Finnish and Swedish-language television programmes shall be accompanied with subtitling or audio subtitling or a service that converts the subtitles of a programme into speech (audio-subtitling and subtitling service) as specified in this section. • The audio-subtitling and subtitling service shall be included in public service programmes as referred to in the Act on Yleisradio Oy. In accordance with further provisions laid down by Government decree, the service shall also be available for programmes that are in the public interest and broadcast under national programming licence. Audio and subtitling service is required for musical performances or sports and children's programmes. The costs incurred by the implementation of the audio and subtitling service to other than public service television broadcasting operators may not exceed one per cent of the operator's turnover achieved in the previous financial year. Gradually increasing quotas, expressed as shares of hours of programs, may be laid down by Government decree for programmes that need to be accompanied with the audio-subtitling and subtitling service. In the case of programmes broadcast by virtue of a national programming licence as referred to above in section 2, the quota in 2011-2016 may total 10–	

	50 per cent of programmes and in the case of public service programmes, 50–100 per cent. Provisions on the technical implementation and transmission of the audio and subtitling service may be issued by Government decree. Costs incurred by the implementation of the audio and subtitling service per one hour of programmes shall be laid down by Government decree for two calendar years at a time.	
France	**Law on Freedom of Communication (2016)** For the blind: • Article 27, adaptation of cinematographic products • Article 28, the services exceeding a 2'5% of share are accessible for blind people or with visual impairment. • Article 33, "This contribution can take into account the adaptation of the product for blind people or with visual impairments" (About the spending of public budget) Private services: • Article 33-1. For television broadcast services that exceed the 2.5% of share from the total of all television services. The agreement concerns also the proportion of programmes that, by means of adapted devices and, particularly, in prime time hours, are accessible for blind people or with visual impairment. • Article 34-2: Any distributor of services makes available to the public freely the services for deaf and hard of hearing people and for blind and visually impaired people related to the programmes of the television services they offer. They are responsible for the necessary technical provisions.	**Chart of Audio Description. Principles and Orientations.** Morisset & Gonant (2008) • What, how, where…, voices to be used for AD (man and woman), rhythm must be adjusted to the emotion of the scene but with normal tone… general recommendations, costs, work-load, dedication, creativity… nothing about AST.
Germany	**State Broadcasting Treaty (2013)** • Section 3 (2) all private and public broadcasting services are supposed to increase their offer of accessible programs (there is no definition of accessibility nor any time line envisaged) and broadcastings within their frame of technical and financial possibilities.	**Wenn aus Bildern Worte werden : durch Audio-Description zum Hörfilm** Dosch & Benecke (2004)
Greece	**Presidential Decree 109/2010 on the harmonisation with Directive 2010/13/EU (2010)** • Article 8 mentions accessibility in general for people with visual and hearing disabilities. Percentages about subtitles and news reports with sign language but nothing about AD and AST.	**Audio description guidelines for Greek – A working document (2008)** by Yota Georgakopoulou • Graphics: For text in subtitles, if it's confusing type the word "subtitle" before the subtitled text and then the text in quotes.
Ireland	**Broadcasting Act (2009)** • Section 38 (4) An annual report shall include a report to the Minister on progress made towards increasing accessibility of broadcasting services to people with disabilities, and in particular, on	**Broadcast Authority of Ireland Guidelines Audio Description** Broadcast Authority of Ireland (2012) • Subtitled captions: any on-screen signs of writing which are relevant

	progress made to achieve the targets set out in any broadcasting rules. • Section 43 (2) Without prejudice to the generality off subsection (1) (c), broadcasting rules with respect to that paragraph shall require each broadcaster of audio-visual material to take specified steps to provide access to that material by persons who are deaf or have a hearing impairment, persons who are blind or partially sighted, and persons who have a hearing impairment and are partially sighted by means of specified services such as … (c) subtitling, and audio description… • Section 100 (ai) The Authority, in advising the Minister on the sectoral impact of a proposal under this part, shall consider the following matters the extent to which the proposal impacts on the availability, choice, quality and accessibility of services for audiences…	should be described. • 9. Current affairs documentaries. Subtitles or captions within such programmes should be described.
Italy	**Service Provision Agreement Rai 2018-2022** • The RAI has to ensure the adoption of article 32, section 6, in the TUSMAR and of the article 30, section 1, letter b, of the UN Convention on the rights of persons with disabilities of 13 December 2016	**Manual for the aspiring audio descriptor of audiofilms for the visually impaired.** Bussarello & Sordo (2011)
Luxembourg	**Law on Electronic Media (2015)** • Encouragement to make contents accessible.	
Latvia	**Electronic Mass Media Law (2010)** • Article 24 (5) The electronic mass media shall develop a publicly available code of conduct where they indicate the basic principles of the operation thereof, the accepted conditions of ethical activity, regulations regarding inappropriate audio and audiovisual commercial communications, including those the target audience of which is minors and which may negatively affect the psychological and physical development of minors, as well as shall indicate the measures which facilitate the availability of services for persons with impaired vision or hearing. • Section 71 The public service remit is the totality of measures the task of which is • 19) to envisage access to certain broadcasts for people with impaired vision and hearing;	
Lithuania	**Law on the Provision of Information to the People (2015)**	
Malta	**Broadcasting Act (2015)** • Chapter 350 of the Malta Law, 16J (3) Media Service providers shall encourage that their services are gradually made accessible to people with a visual or hearing disability.	

Netherlands	**Media Decision (2010)** • Article 15 of the, only subtitles for the deaf or hard of hearing.	
Poland	**Act on Radio and Television (2018)** • Thresholds for access services at 15% of broadcast per quarter in 2019, 25% in 2020-2021, 35% in 2022-2023 and 50% from 2024 on	**Standardy Tworzenia Audiodeskrypcji do Produkcji Audiowisualnych** Audiodeskrypcji Fundation (2010) **Audiodeskrypcja – zasady tworzenia** Fundacja Kultury bez Barier (2012)
Portugal	**Multiannual plan that defines the set of obligations concerning accessibility to television programmes and audiovisual services for people with disabilities (2016)** • Public channels - RTP1: 70 hours of AD programmes a year - RTP2: 20 hours of AD programmes a year • Private channels SIC and TVI: 12 hours of AD programmes a year.	
Romania	**Romanian Audio-visual Law (2002)** • The National Council of the Audio-Visual has the obligation to encourage the providers of audio-visual services in view of ensuring conditions so that the services provided become accessible to persons with visual and hearing impairments. • Romanian audio-visual code adopted by the National Council of the Audio-visual, "having in view of the need to provide audio-visual media accessible for the persons with a disability and the elderly, though sign language, subtitles, audio-description and other technical means offered by digital technology"	
Slovakia	**Broadcasting and Retransmission that amended Law on Telecommunications (2000)** • Article 18. A public service broadcaster shall have the duty to ensure multimodal access to its television programme service, such that in every television programme service that it broadcasts digitally, are at least: c) 20% of all broadcast programmes accompanied by a voice commentary for the blind. • Article 18a A licenced broadcaster shall have the duty to ensure multimodal access to its programme service such that in every television programme service that it broadcasts digitally, are at least b) 3% of all broadcast programme accompanied by a voice commentary for the blind. • Article 18b Specific Obligations of Broadcasters and Providers of an On-demand Audiovisual Media Service to Mark Programmes with	

	Multimodal Access (1), broadcasters and providers of an on-demand audiovisual media service shall clearly mark all programmes that are accompanied by hidden or displayed subtitles that correspond to the plot of broadcast programmes, are accompanied voice commentary for the blind.	
Slovenia	**Audiovisual Media Services Act (2011)** • Article 11 in accordance with the law governing the media, encourages providers to gradually ensure access to their services for persons with visual and hearing impairment.	
Spain	**Law on General Audiovisual Communication (2015)** • Article 6.3. All information referred to in this article found in the webpages, electronic guides of programs and other communication means of the providers of audiovisual communication services that make effective the right to transparency regulated in the present article, have to be made accessible to people with disabilities. • Article 8 People with visual or hearing impairment have the right to a universal accessibility to audiovisual communication considering the technological possibilities. • Article 8.3. People with visual impairment have the right to two audio described hours a week of the television audiovisual communication, free and national or regional coverage. About transitional laws, No. 5. • The accessibility services for people with disabilities in the programming of the channels found in article 8 will have to reach on December 31st each year the following percentages and values: For subtitling: 25% in 2010, 45% in 3011, 65% in 2012 and 75% in 2013; for sign language and audio description (in hours): 0.5 in 2010, 1 in 2011, 1.5 in 2012 and 2 in 2013 • The accessibility services for persons with disability to be offered as part of public programming should at least have reached the following percentages and values by December 31st of each year: For subtitling: 25% in 2010, 50%in 2011, 70% in 2012 and 90% in 2013; for sign language and audio description (in hours): 1 in 2010, 3 in 2011, 7 in 2012 and 10 in 2013. **Law 10/2005 on Urgent Measures for the stimulation of Digital Terrestrial Television, of the Liberation of Cable Television and the Encouragement of Pluralism.**	**Audio description for people with visual impairment. Requirements for the audio description and creation of audio guides.** AENOR (2005) • In the script the information provided by occasional subtitles, warnings and credits must be included by summarising those that are excessively long when the space for the message is short to allow for a literal audio description.

	• "Pursuant to the provisions of Act 51/2003, of 2 December, on equal opportunities, non-discrimination and universal accessibility for persons with disabilities, and to its rules of implementation, the competent administrations shall adopt, after having consulted the representatives of the sectors and persons concerned, the necessary measures to guarantee from the start accessibility of persons with disability to the services of digital terrestrial television. To this purpose, the adopted measures will respect the principles of universal accessibility and design for all persons concerned".	
Sweden	**Proposed Government Bill Education and Accessibility – Public Service Broadcasting 2014-2019 (2014)** • Contains conditions and guidelines for 3 channels (Sveriges Radio AG (SR), Sveriges Television AB (SVT) and Sveriges Utbildningsradio AB (UR). • The proportion of programmes with sign language interpretation and audio description is to increase during the licence period. • The spoken text service should be offered in all programmes that are not live broadcast.	
United Kingdom	**Equality Act (2010)** • Only BBC and Channel 4 are public authorities and subjected to the PSED (Public Sector Equality Duty). Laws only for public media providers. • BBC's Telling our story. Equality and diversity at the BBC 2011. It provides BBC's performance for the period 2010-2011 with Audio description in 12% of their programs (we don't know about AST). Also they aim to make more services available for all kind of people with disabilities both online and offline. • Channel 4's webpage, about their commitment. According to Ofcom codes. 2016. Audio description: 10% on Channel 4, E4, More4, Film4 and 7.7% on 4Seven.	**Code on Television Access Services** Office of Communications (2017) • About specific recommendations for the issuance of AD it recommends ITC guidelines • Best practices: about what to describe (and it says on-screen information I guess concerning AST as well), description of characters, on-screen action, settings, what not to describe, when to describe, language, delivery, balance. • A1.30. […] To differentiate between subtitles and description the describer should do this by either the use of their voice (e.g. stating the obvious, 'He says in Russian...' or 'A caption reads...') or a second voice. **Audio Description Guidelines** O'Hara, n.d. What to describe: • Subtitled captions • Points to consider: • How does the describer differentiate between subtitles and description? By use of the voice, or by stating the obvious: *'He says in Russian…'*? Or simply: *"A caption reads…"*?

United States of America	**Twenty-First Century Communications and Video Accessibility (2010) and Video Description: Implementation of the Twenty-Frist Century Communications and video Accessibility Act of 2010 (2016)** Title II – Video Programming • Requires video programming distributors, providers, and owners to convey emergency information in a manner that is accessible to people who are blind or visually impaired. • Expands the requirement for video programming equipment (equipment that shows TV programs) to be capable of displaying closed captions, to devices with screens smaller than 13 inches (e.g., portable TVs, laptops, smart phones), and requires these devices to be able to pass through video descriptions and emergency information that is accessible to people who are blind or visually impaired, if technically feasible and achievable. **Rehabilitation Act (1973)** • Section 508, new amendment 1998 expanding the act to include equal access to electronic and information technology. • Section 504, it also implicates federally funded programs. **NY State Accessibility Policy (2010)** • For state websites. A. 11.2 – Video/Visual: State agencies will provide a video description for multimedia content that contains video, or other visual information necessary for the understanding of the content.	**Audio description guidelines and best practices American Council of the Blind's Audio Description Project.** Snyder (2010) • Reading disclaimers and credits at the beginning and end of films, videos and television programs is an important function of audio description. As well, the describer should read text and subtitles. • … introduce text or subtitles with a phrase such as, "Words appear" or "Subtitles appear". About surtitles for opera: • Normally two people, man and woman, to read translation/to describe • Change of tone to voice surtitles to differentiate character • Advanced copy of the surtitles is extremely helpful in this regard. **Audio Description Style Guide** Netflix (n.d.) • 2.0 Describing On-screen Elements • 2.2. Subtitles for foreign language • The description should introduce subtitles by adjusting the tone of voice for distinction and read the subtitles verbatim. The original dialogue audio should be dipped in order to avoid confusion, but still allow the viewer to hear the original dialogue in the background. State "subtitles" when necessary to avoid confusion
International		**Technical Report: Part 5: Final report of activities: Working Group B "Audio/Video description and spoken captions"** International Telecommunication Union (2013) • Final report activities: "Audio/Video description and spoken captions" **2.2 Differences in practice worldwide** • One additional issue is related to whether AD is destined for foreign language productions or multilingual productions in what are traditionally known as "dubbing countries" or "subtitling countries", even though this distinction is no longer water tight. Dubbed films can integrate AD

		in the language of the dubbed dialogues. Subtitled films must provide an aural rendering of the subtitles, or audio subtitles, since the target audience will otherwise hear an AD in language B and hear film dialogues in language A. The translation in the form of written subtitles that appear as text on screen is obviously not accessible to the core target audience. • Much research is still required into the best practice for AST with regard to: • Conditions under which AST does and does not work (e.g. non-fiction interviews lend themselves well to AST, film dialogues may not); • The degree to which subtitles can (copyright) and should be rewritten to be spoken and interact with AD (subtitles may leave out essential information); • The type of recording that works best (voice-over Internet Protocol (VoIP), dubbing without lip sync or other) the number and types of voices to be used for AST; • The feasibility of using text-to-speech for AST as well as AD; • AST does not replace AD, since it only deals with existing dialogues, and AD will still be needed to make the production accessible. **2.3. Uneven legislative requirements across countries and lack of uniform definitions of blindness** • This has an immediate impact on what services are made available for these ill-defined target groups, what is subsidised and on what conditions, including AD and AST, and AD+AST. **3. What is the vision for 2015-2020** **3.1. The digital switch** • Integrated broadcast and broadband service also offers great possibilities to users who will be able to customize the service to their personal needs. What is also important is that it allows for the focused introduction of AD and AST, i.e. only to be activated by the target audience when and where they need it.

2. Documents Related

2.1. Phase 1: Online Validation

Information Sheet for the Online Validation

HOJA INFORMATIVA SOBRE EL EXPERIMENTO

El objetivo general del proyecto es validar una serie de escenas de películas para entender qué emociones suscitan.

Vuestra participación en el experimento será la siguiente: tenéis que ver diversas escenas de diferentes películas.

Después de cada clip tendréis que completar unas preguntas siguiendo las instrucciones que os daremos.

Vuestra participación es totalmente voluntaria y os podéis retirar del estudio en cualquier momento sin tenerlo que justificar de ninguna manera y sin que esto os afecte negativamente de cualquier modo.

Vuestros datos serán totalmente anónimos ya que en el formulario online no se precisa ninguna información personal.

Los investigadores que llevarán a cabo el experimento son Gonzalo Iturregui Gallardo (que participará en las pruebas) y la Dra. Anna Matamala (que dirige la investigación). Esta última es la responsable y la podéis contactar en anna.matamala@uab.cat, en el despacho K-2012 de la UAB o bien llamando a 935813360.

Si queréis más información sobre las evoluciones de la investigación, tan sólo hace falta pedirlo a la investigadora principal por correo electrónico.

¡MUCHAS GRACIAS por vuestra participación!

(Nota: se le facilita también el documento de consentimiento informado)

Consent Form for the Online Validation

CONSENTIMIENTO INFORMADO

Departamento de Traducción, Interpretación y Estudios de Asia Oriental
Universidad Autónoma de Barcelona

Nombre del proyecto:

Evaluación de la activación emocional en diferentes estrategias de audiosubtitulación (Validación emocional online de escenas cinematográficas)

- He leído la hoja informativa que me han dado así como el consentimiento informado.
- He recibido suficiente información sobre el estudio y la he entendido.

Por lo tanto,

- Entiendo que mi participación es voluntaria.
- Entiendo que mi información será confidencial.
- Entiendo que puedo retirarme del estudio cuando quiera y sin tener que dar explicaciones y sin que haya repercusiones negativas.

Haciendo clic sobre el botón de sí, doy libremente mi conformidad para poder participar en el estudio.

2.2. Phase 2: T-SAM Validation (Focus Group)

Protocol for T-SAM Validation

NEA	**Validació d'una versió tàctil del qüestionari SAM**
Finançat pel Ministeri d'Economia, Indústria i Competitivitat ((FFI2015-64038-P, MINECO/FEDER, UE) dins del grup de recerca Transmedia Catalonia (2014SGR0027)	

VALIDACIÓ D'UNA VERSIÓ TÀCTIL DEL QÜESTIONARI SAM

GRUP FOCAL

Introducció (5 minuts)

Donar la benvinguda, presentar l'objectiu i el context del grup focal, explicar què és un grup focal i com es desenvolupa, i fer presentacions.

Objectiu: **Validació d'una versió tàctil del qüestionari SAM**

Dades demogràfiques (5 minuts)

Edat, gènere i nivell de dèficit visual.

Tasca 1. Explicació i comprensió del qüestionari SAM (10 minuts).

El qüestionari SAM, *Self-Assessment Mannequin* [Modelo de Autoevaluación], va ser creat per Bradley i Lang (1994) per a mesurar l'activació emocional.

El qüestionari està dividit en dos franges de 5 imatges cada una, disposades una al costat de l'altra. Les cinc imatges segueixen una progressió que es desenvolupa d'esquerra a dreta.

Els dos qüestionaris que es presenten als participants són versions tàctils del mateix qüestionari. Ambdós es basen en les mateixes idees gràfiques però compten amb petites diferències que poden millorar la seva interpretació.

A la primera franja, el participant ha de valorar l'emoció de negativa a positiva. L'avatar que apareix a les imatges, que es tracta d'una cara humana, representa una cara trista amb els llavis corbats cap a baix i amb les celles baixes. La cara esdevé més i més contenta, amb els llavis que es corben progressivament cap amunt, passant per una cara en repòs fins a arribar a una cara contenta amb les celles cap amunt.

A la segona franja, les cinc imatges expressen com d'activat estàs per l'emoció. Mostra cinc cercles dels quals la línia es deforma cap a marges més ondulats fins que explota en una figura punxeguda i afilada, com si fos la representació d'una explosió en un còmic.

Sota cadascuna de les franges d'imatges, apareix una sèrie de números de l'1 al 9. Així el número 1 correspons a la primera imatge, el número 2 queda entre la primera i la segona imatge; el número 3 correspons a la segona imatge, el número 4 queda entre la segona i la tercera imatge, i així successivament.

Preguntes

Tasca 2. Funcionament del qüestionari.

1. Has entès el principal objectiu del qüestionari?

2. Quins conceptes valoren la primera i la segona franges del qüestionari?

Tasca 3. Primera franja.

1. Entens la representació gràfica d'una cara humana?

    1.1. En quin dels dos és més clara?

2. És difícil d'entendre per mitja del tacte la forma dels llavis? I la de les celles?

    1.2. En quin dels dos és més clar?

3. Estàs familiaritzat amb els emoticones i maneres similars de representar les expressions humanes?

4. Ajuden a entendre la cara humana el contorn de la cara i la inclusió del nas i dels ulls? Preferiries que no hi fossin?

Tasca 4. Segona franja.

1. Reconeixes la forma de les figures representades a les imatges?

    1.2. En quin dels dos qüestionaris queda més clar?

2. Relaciones les línies de les formes amb l'activació emocional? Suau i recte vol dir calmat i punxegut i afilat vol dir excitat?

3. Se t'acudeixen altres formes de representar-ho?

Tasca 5. Resum genera i discussió

1. Diries que el qüestionari es pot completar sense cap altra ajuda a part de l'explicació inicial?

2. En general, quin diries que és el qüestionari més fàcil d'entendre?

3. Quines característiques d'un o de l'altre emfatitzaries com a estratègies que poden facilitar la seva lectura i interpretació?

Consent Form for the T-SAM Validation

CONSENTIMIENTO INFORMADO

Departamento de Traducción, Interpretación y Estudios de Asia Oriental
Universidad Autónoma de Barcelona

Nombre del proyecto:

Evaluación de la activación emocional en diferentes estrategias de audiosubtitulación
Validación de la versión táctil del cuestionario SAM

Nombre de los investigadores: Gonzalo Iturregui Gallardo (gonzalo.iturregui@uab.cat) y Anna Matamala (anna.matamala@uab.cat).

¿Podría indicarnos su nombre y apellidos? __________________________________

¿Ha recibido la suficiente información sobre el grupo focal y la ha entendido? Por favor, sí o no. _____

¿Ha entendido el consentimiento informado? Por favor, responda sí o no. _____

¿Ha podido hacer preguntas sobre el grupo focal? Por favor, responda sí o no. _____

¿Ha hablado con Gonzalo Iturregui Gallardo? Por favor, responda sí o no. _____

¿Ha entendido que su participación es voluntaria? Por favor, responda sí o no. _____

¿Ha entendido que su información será confidencial? Por favor, responda sí o no. _____

¿Ha entendido que puede retirarse del estudio cuando quiera y sin tener que dar explicaciones y sin que haya repercusiones negativas? Por favor, responda sí o no. _____

¿Da usted libremente su conformidad para poder participar en el estudio? Por favor, responda sí o no. _____

¿Da usted permiso para que se grave la sesión de grupo focal? Por favor, responda sí o no. _____

¿El investigador, Gonzalo Iturregui Gallardo, asegura haber informado al participante debidamente? Responda sí o no. _____

Barcelona, ____ de _____________________ de 2018

2.3. Phase 2: Experiment 1

Information Sheet for Experiment 1

HOJA INFORMATIVA SOBRE EL EXPERIMENTO

El objetivo general del proyecto es observar la recepción de los usuarios al ver diferentes películas extranjeras subtituladas/audiosubtituladas.

Vuestra participación en el experimento será la siguiente: tenéis que ver 4 escenas de diferentes películas.

Antes de empezar el experimento os pediremos que os coloquéis dos sensores para obtener datos del batido del corazón y la conductividad de la piel. Os ayudaremos si es necesario. El primer sensor es un cinturón que se coloca justo debajo del pecho, como los que llevan algunos deportistas para medir los latidos del corazón; y el segundo está formado por dos pequeños receptores que se sitúan bajo la segunda falange de los dedos índice y anular.

Después de cada clip tendréis que completar los cuestionarios siguiendo las instrucciones que os daremos. Dispondremos también de cuestionarios en formato electrónico para aquellas personas que lo precisen.

Vuestra participación es totalmente voluntaria y os podéis retirar del estudio en cualquier momento sin tenerlo que justificar de ninguna manera y sin que esto os afecte negativamente de cualquier modo.

Vuestros datos serán totalmente anónimos ya que en la hoja escrita sólo os pediremos que proporcionéis vuestro sexo (hombre/mujer), edad, vuestros hábitos audiovisuales y la profesión.

Los investigadores que llevarán a cabo el experimento son Gonzalo Iturregui Gallardo (que participará en las pruebas) y la Dra. Anna Matamala (que dirige la investigación). Esta última es la responsable y la podéis contactar en anna.matamala@uab.cat, en el despacho K-2012 de la UAB o bien llamando a 935813360.

Si queréis más información sobre las evoluciones de la investigación, tan sólo hace falta pedirlo a la investigadora principal por correo electrónico.

¡MUCHAS GRACIAS por vuestra participación!

(Nota: se les facilita también el documento de consentimiento informado de forma oral y escrita)

Consent Form for Experiment 1

CONSENTIMIENTO INFORMADO

Departamento de Traducción, Interpretación y Estudios de Asia Oriental
Universidad Autónoma de Barcelona

Nombre del proyecto:

Evaluación de la activación emocional en diferentes estrategias de audiosubtitulación (Subexperimento 1)

Nombre de los investigadores: Gonzalo Iturregui Gallardo (gonzalo.iturregui@uab.cat) y Anna Matamala (anna.matamala@uab.cat).

¿Podría indicar-nos su nombre y apellidos? ___________________________________

- ¿Ha entendido el contenido de la hoja informativa y del consentimiento informado? Por favor, responda sí o no. ______________
- ¿Ha recibido la suficiente información sobre el estudio y la ha entendido? Por favor, responda sí o no. ____________
- ¿Ha podido hacer preguntas sobre el estudio? Por favor, responda sí o no. ___________

¿Ha hablado con Gonzalo Iturregui Gallardo? Por favor, responda sí o no. _______________

- ¿Ha entendido que su participación es voluntaria? Por favor, responda sí o no. _________
- ¿Ha entendido que su información será confidencial? Por favor, responda sí o no. ___________
- ¿Ha entendido que puede retirarse del estudio cuando quiera y sin tener que dar explicaciones y sin que haya repercusiones negativas? Por favor, responda sí o no. ___________

¿Da usted libremente su conformidad para poder participar en el estudio? Por favor, responda sí o no. _____________

¿El investigador, Gonzalo Iturregui Gallardo, asegura haber informado al participante debidamente? Responda sí o no. __________

Pre-Questionnaire: Experiment 1 (Spanish)

Muchas gracias por participar en este experimento.
A continuación se le realizarán algunas preguntas que nos ayudarán a mejorar los resultados de este estudio.
Por favor, responda lo más sinceramente posible y recuerde que este cuestionario es anónimo.

Código de participante

____

Cuestiones previas

1. ¿Género?

- o Masculino
- o Femenino
- o Prefiero no contestar
- o Otro

2. Edad

____

3. Me considero una persona…

- o Ciega
- o Con baja visión

4. Mi discapacidad visual empezó…

- o De nacimiento
- o Entre los 0 y los 4 años
- o Entre los 5 y los 12 años
- o Entre los 13 y los 20 años
- o Entre los 21 y los 40 años
- o Entre los 40 y los 60 años
- o Con más de 60 años

5. Estudios finalizados

- o Sin estudios
- o Educación primaria
- o Educación secundaria
- o Formación especializada
- o Universidad

6. ¿Está en la actualidad consumiendo algún tipo de medicación para la ansiedad?

- o Sí
- o No

En caso afirmativo, ¿cuál?

____

7. ¿Ha consumido cafeína en las últimas dos horas?

- o Sí
- o No

8. ¿Ha consumido alcohol en las últimas dos horas?

- o Sí
- o No

9. ¿Ha consumido alguna otra droga en las últimas dos horas?

- o Sí
- o No

En caso afirmativo, ¿cuál?

10. ¿Cuántas horas ha dormido aproximadamente en la última noche?

- o Menos de 3 horas
- o De 3 a 6 horas
- o De 6 a 8 horas
- o De 8 a 10 horas
- o Más de 10 horas

11. ¿Ha estado en las últimas dos horas en alguna situación especialmente estresante, como por ejemplo un examen oral?

- o Sí
- o No

En caso afirmativo, ¿cuál?

____

12. ¿Tiene algún tipo de alteración cardiovascular?

- o Sí
- o No

En caso afirmativo, ¿toma alguna medicación para la misma?

____

12. ¿Desea añadir algún comentario?

____

Pre-Questionnaire: Experiment 1 (English)

Thank you for taking part in this experiment.
Now, some questions will be asked in order to improve the results of this study.
Please, answer as sincerely as possible and keep in mind that this questionnaire is anonymous.

Participant code

____

Previous questions

1. ¿Gender?

- o Masculine
- o Feminine
- o No answer
- o Other

2. Age

____

3. I consider myself as…

- o Blind
- o Partially sighted

4. My visual disability started…

- o From birth
- o Between 0 and 4 years old
- o Between 5 and 12 years old
- o Between 13 and 20 years old
- o Between 21 and 40 years old
- o Between 40 and 60 years old
- o With moret han 60 years old

5. Education level

- o No education
- o Primary education
- o Secondary education
- o Specialised training
- o University

6. Are you currently taking any medicine for the anxiety?

- o Yes
- o No

If yes, which one?

____

7. Have you consumed cafeine in the last two hours?

- Yes
- No

8. Have you consumed alcohol in the last two hours?

- Yes
- No

9. Have you consumed any other drug in the last two hours?

- Yes
- No

If yes, which one?

10. How many hours did you sleep last night?

- Less than 3 hours
- From 3 to 6 hours
- From 6 to 8 hours
- From 8 to 10 hours
- More than 10 hours

11. Have you been in any particularly stressing situation in the last two hours, such as an oral exam?

- Yes
- No

If yes, which one?

____

12. Do you have any kind of cardiovascular condition?

- Yes
- No

If yes, do you take any medicine to treat it?

____

12. Do you have any other comment?

____

Preference Questionnaire

Audio Subtitling

1. Which of the two options do you prefer?

- The more acted voice in which there is no original.
- The less acted voice in which the original can be heard.
- Irrelevant.

Why?

2. Which of the two options do you think can transmit better the emotion of the scene?

- The more acted voice in which there is no original.
- The less acted voice in which the original can be heard.
- Both options transmit the same emotion.

Why?

3. Was it more difficult to follow the dialogues in any of the two options? Which one?

- The more acted voice in which there is no original.
- The less acted voice in which the original can be heard.
- I followed the dialogues in the same way.

Why?

4. Which voice do you prefer for the audio subtitles (when they are read aloud)?

- Acted voice.
- Less acted voice/read aloud.
- Irrelevant.

Why?

Two voices were used to deliver the subtitles: a female voice for all female characters and a male voice for all male characters.

Would you rather have a different voice for each character?

- Yes.
- No.
- Irrelevant.

Why?

6. How would you like the audio subtitle's voice to be?

Audio Description

The voice of the audio description was the same for both versions. The intonation used was quite neutral.

7. Would you rather have a more acted audio description?

- Yes.
- No.
- Irrelevant.

Why?

8. How do you like the audio description's voice to be?

2.4. Phase 2: Experiment 2

Information Sheet for Experiment 2

HOJA INFORMATIVA SOBRE EL EXPERIMENTO

El objetivo general del proyecto es observar la recepción de los usuarios al ver diferentes películas extranjeras subtituladas/audiosubtituladas.

Vuestra participación en el experimento será la siguiente: tenéis que ver 4 escenas de diferentes películas.

Antes de empezar el experimento os pediremos que os coloquéis dos sensores para obtener datos del batido del corazón y la conductividad de la piel. Os ayudaremos si es necesario. El primer sensor es un cinturón que se coloca justo debajo del pecho, como los que llevan algunos deportistas para medir los latidos del corazón; y el segundo está formado por dos pequeños receptores que se sitúan bajo la segunda falange de los dedos índice y anular.

Después de cada clip tendréis que completar los cuestionarios siguiendo las instrucciones que os daremos. Dispondremos también de cuestionarios en formato electrónico para aquellas personas que lo precisen.

Vuestra participación es totalmente voluntaria y os podéis retirar del estudio en cualquier momento sin tenerlo que justificar de ninguna manera y sin que esto os afecte negativamente de cualquier modo.

Vuestros datos serán totalmente anónimos ya que en la hoja escrita sólo os pediremos que proporcionéis vuestro sexo (hombre/mujer), edad, vuestros hábitos audiovisuales y la profesión.

Los investigadores que llevarán a cabo el experimento son Gonzalo Iturregui Gallardo (que participará en las pruebas) y la Dra. Anna Matamala (que dirige la investigación). Esta última es la responsable y la podéis contactar en anna.matamala@uab.cat, en el despacho K-2012 de la UAB o bien llamando a 935813360.

Si queréis más información sobre las evoluciones de la investigación, tan sólo hace falta pedirlo a la investigadora principal por correo electrónico.

¡MUCHAS GRACIAS por vuestra participación!

(Nota: se les facilita también el documento de consentimiento informado)

Consent Form for Experiment 2

CONSENTIMIENTO INFORMADO

Departamento de Traducción, Interpretación y Estudios de Asia Oriental
Universidad Autónoma de Barcelona

Nombre del proyecto:

Evaluación de la activación emocional en diferentes estrategias de audiosubtitulación (Subexperimento 2)

Nombre del investigador: Gonzalo Iturregui Gallardo (gonzalo.iturregui@uab.cat) y Anna Matamala (anna.matamala@uab.cat).

Yo, (NOMBRE Y APELLIDOS) _________________________________

- He leído la hoja informativa que me han dado así como el consentimiento informado.
- He recibido suficiente información sobre el estudio y la he entendido.
- He podido hacer preguntas sobre el estudio.

He hablado con: Gonzalo Iturregui Gallardo.

- Entiendo que mi participación es voluntaria.
- Entiendo que mi información será confidencial.
- Entiendo que puedo retirarme del estudio cuando quiera y sin tener que dar explicaciones y sin que haya repercusiones negativas.

Doy libremente mi conformidad para poder participar en el estudio.

Firma del investigador Firma del participante

Gonzalo Iturregui Gallardo

(Nota: en la otra cara de esta hoja encontrareis el documento informativo)

Pre-Questionnaire: Experiment 2 (Spanish)

Muchas gracias por participar en este experimento.
A continuación se le realizarán algunas preguntas que nos ayudarán a mejorar los resultados de este estudio.
Por favor, responda lo más sinceramente posible y recuerde que este cuestionario es anónimo.

Código de participante

____

Cuestiones previas

1. ¿Género?

- o Masculino
- o Femenino
- o Prefiero no contestar
- o Otro

2. Edad

____

- o

3. Estudios finalizados

- o Sin estudios
- o Educación primaria
- o Educación secundaria
- o Formación especializada
- o Universidad

4. ¿Está en la actualidad consumiendo algún tipo de medicación para la ansiedad?

- o Sí
- o No

En caso afirmativo, ¿cuál?

____

5. ¿Ha consumido cafeína en las últimas dos horas?

- o Sí
- o No

6. ¿Ha consumido alcohol en las últimas dos horas?

- o Sí
- o No

7. ¿Ha consumido alguna otra droga en las últimas dos horas?

- Sí
- No

En caso afirmativo, ¿cuál?

8. ¿Cuántas horas ha dormido aproximadamente en la última noche?

- Menos de 3 horas
- De 3 a 6 horas
- De 6 a 8 horas
- De 8 a 10 horas
- Más de 10 horas

9. ¿Ha estado en las últimas dos horas en alguna situación especialmente estresante, como por ejemplo un examen oral?

- Sí
- No

En caso afirmativo, ¿cuál?
____

10. ¿Tiene algún tipo de alteración cardiovascular?

- Sí
- No

En caso afirmativo, ¿toma alguna medicación para la misma?
____

11. ¿Desea añadir algún comentario?
____

Pre-Questionnaire: Experiment 2 (English)

Thank you for taking part in this experiment.
Now, some questions will be asked in order to improve the results of this study.
Please, answer as sincerely as possible and keep in mind that this questionnaire is anonymous.

Participant code

____

Previous questions

1. ¿Gender?

- Masculine
- Feminine
- No answer
- Other

2. Age

____

3. Education level

- No education
- Primary education
- Secondary education
- Specialised training
- University

4. Are you currently taking any medicine for the anxiety?

- Yes
- No

If yes, which one?

____

5. Have you consumed cafeine in the last two hours?

- Yes
- No

6. Have you consumed alcohol in the last two hours?

- Yes
- No

7. Have you consumed any other drug in the last two hours?

- Yes
- No

If yes, which one?

8. How many hours did you sleep last night?

- o Less than 3 hours
- o From 3 to 6 hours
- o From 6 to 8 hours
- o From 8 to 10 hours
- o More than 10 hours

9. Have you been in any particularly stressing situation in the last two hours, such as an oral exam?

- o Yes
- o No

If yes, which one?

____

10. Do you have any kind of cardiovascular condition?

- o Yes
- o No

If yes, do you take any medicine to treat it?

____

11. Do you have any other comment?

____

Alexithymia Questionnaire (TAS-20) (Spanish)

Cuestionario adaptado TAS-20: escala de Alexitimia de Toronto de 20 elementos.
De las características que se le van a presentar a continuación, señale el grado en que se ajusten a su modo de ser habitual. Por favor, conteste lo mas sinceramente posible.

Valores:
1: Muy en desacuerdo
2: En desacuerdo
3: Ligeramente en desacuerdo
4: Ligeramente de acuerdo
5: De acuerdo
6: Muy de acuerdo

1. A menudo me encuentro confundido sobre cuál es la emoción que estoy sintiendo.

1	2	3	4	5	6
○	○	○	○	○	○

2. Me es difícil encontrar las palabras exactas para mis emociones.

1	2	3	4	5	6
○	○	○	○	○	○

3. Tengo sensaciones físicas que ni los médicos entienden.

1	2	3	4	5	6
○	○	○	○	○	○

4. Soy capaz de describir mis sentimientos con facilidad.

1	2	3	4	5	6
○	○	○	○	○	○

5. Prefiero analizar los problemas en vez de describirlos.

1	2	3	4	5	6
○	○	○	○	○	○

6. Cuando estoy mal no se si estoy triste, asustado o enfadado.

1	2	3	4	5	6
○	○	○	○	○	○

7. A menudo estoy hecho un lío con las sensaciones que noto en mi cuerpo.

1	2	3	4	5	6
○	○	○	○	○	○

8. Prefiero dejar que pasen las cosas, en vez de analizar por qué han ocurrido así.

1	2	3	4	5	6
○	○	○	○	○	○

9. Tengo sentimientos que no puedo identificar.

1	2	3	4	5	6
○	○	○	○	○	○

10. Es esencial estar en contacto con la gente y sentir emociones y sentimientos.

1	2	3	4	5	6
○	○	○	○	○	○

11. Me es difícil describir o explicar lo que siento sobre la gente.

1	2	3	4	5	6
○	○	○	○	○	○

12. La gente me pide que explique con más detalle mis sentimientos.

1	2	3	4	5	6
○	○	○	○	○	○

13. No sé lo que está pasando dentro de mí.

1	2	3	4	5	6
○	○	○	○	○	○

14. A menudo no sé por qué estoy enfadado.

1	2	3	4	5	6
○	○	○	○	○	○

15. Prefiero hablar con la gente de sus actividades diarias más que de sus sentimientos.

1	2	3	4	5	6
○	○	○	○	○	○

16. Prefiero ver programas de TV de entretenimiento, superficiales, en vez de dramas psicológicos.

1	2	3	4	5	6
○	○	○	○	○	○

17. Me es difícil revelar mis más profundos sentimientos, incluso a mis amigos íntimos.

1	2	3	4	5	6
○	○	○	○	○	○

18. Puedo sentirme cercano a alguien, comprender sus sentimientos, incluso en momentos de silencio.

1	2	3	4	5	6
○	○	○	○	○	○

19. Me resulta útil examinar mis sentimientos para resolver problemas personales.

1	2	3	4	5	6
○	○	○	○	○	○

20. Buscar el significado oculto de las películas distrae de la diversión que se pueda sentir viéndolas simplemente.

1	2	3	4	5	6
○	○	○	○	○	○

Alexithymia Questionnaire (TAS-20) (English)

Adapted questionnaire TAS-20: Toronto Alexithymia Scale of 20 elements.
In each of the characteristics that are here presented, answer with the degree that most relates more to your personality. Please, answer as sincerely as possible.

Values:
1: Completely disagree
2: Disagree
3: Slightly disagree
4: Slightly agree
5: Agree
6: Completely agree

1. I am often confused about what emotion I am feeling.

1	2	3	4	5	6
○	○	○	○	○	○

2. It is difficult for me to find the right words for my feelings.

1	2	3	4	5	6
○	○	○	○	○	○

3. I have physical sensations that even doctors don’t understand.

1	2	3	4	5	6
○	○	○	○	○	○

4. I am able to describe my feelings easily.

1	2	3	4	5	6
○	○	○	○	○	○

5. I prefer to analyse problems rather tan just describe them.

1	2	3	4	5	6
○	○	○	○	○	○

6. When I am upset, I don’t know if I am sad, frightened, or angry.

1	2	3	4	5	6
○	○	○	○	○	○

7. I am often puzzled by sensations in my body.

1	2	3	4	5	6
○	○	○	○	○	○

8. I prefer to just let things happen rather than to understand why they turned out that way.

1 2 3 4 5 6

9. I have feelings that I can't quite identify..

1 2 3 4 5 6

10. Being in touch with emotions is essential..

1 2 3 4 5 6

11. I find it hard to describe how I feel about people.

1 2 3 4 5 6

12. People tell me to describe my feelings more.

1 2 3 4 5 6

13. I don't know what's going on inside me.

1 2 3 4 5 6

14. I often don't know why I am angry.

1 2 3 4 5 6

15. I prefer talking to people about their daily activities rather than their feelings.

1 2 3 4 5 6

16. I prefer to watch light entertainment shows rather than psychological dramas.

1 2 3 4 5 6

17. It is difficult for me to reveal my innermost feelings, even to close friends.

1 2 3 4 5 6

18. I can feel close to someone, even in moments of silence.

1 2 3 4 5 6

19. I find examination of my feelings useful in solving personal problems.

1	2	3	4	5	6
○	○	○	○	○	○

20. Looking for hidden meanings in movies distracts from their enjoyment.

1	2	3	4	5	6
○	○	○	○	○	○

Universitat Autònoma
de Barcelona