



Universitat Autònoma de Barcelona

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>

Universitat Autònoma de Barcelona

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

Doctorat en Traducció i Estudis Interculturals

Audio Description in China:

Past, Present and Future

Author:

Irene Tor-Carroggio

Supervised by:

Dr. Helena Casas-Tost

Dr. Sara Rovira-Esteva

2020

To my brother Carlos.

Because there are many ways of becoming a doctor.

Abstract

Despite China being the country with most disabled people in the world, the provision of media accessibility services and their academic study are still in their infancy, including audio description. Audio description is a “description of visual information delivered via an audio channel” (Salway, 2007, p. 151) catering people with sight loss, but not exclusively. Audio description is currently mainly offered in the so-called “first-tier cities” in China, such as Beijing, Shanghai and Guangzhou. This is not surprising considering that the provision of audio description relies on the priorities of each regional government, along with grassroot initiatives. The academic side of audio description does not come off better, although the state of Audio-visual Translation is quite vibrant in China. In the light of the precarious state of audio description and the challenges it still has to face in China, the present thesis sets out to not only contextualise audio description in the country as a first step to launch academic research on the topic, but also offers an audio description-related alternative that will facilitate the provision of audio description. More specifically, the general aims of this thesis are the following. First, to describe the history and the current state of audio description in Mainland China, both as an access service and as an object of study, with special focus on the city of Shanghai. This is carried out by reviewing the history and the current state of audio description in China; profiling audio describers in Mainland China; investigating the habits, needs and satisfaction of audio description users towards this access service; and analysing some Chinese audio description guidelines currently in use. Second, to compare the Chinese case with the European context regarding the aspects mentioned in the first objective. Third, to determine whether text-to-speech audio description can be accepted in movies in China. The methodology employed includes questionnaires, interviews and a focus group, all of which were distributed or carried out in China. Both quantitative and qualitative data were collected. The results reveal that audio description in Mainland China is still in its infancy, both as professional practice and as an object of study, especially when compared to some countries in the West. It is mainly offered in films and thanks to volunteers, who write the scripts and voice them. Our user participants were mainly retired old people with little or no education that consume audio described films quite frequently mainly for entertainment purposes. They are all satisfied with the service provided so far, despite it not being delivered by persons who have been professionally trained in this field. Regarding our reception study, the results show that, although natural voices have

statistically higher scores than synthetic voices, these can be viewed as both an interim alternative and even as a permanent solution by our participants if that translates into more access to audio described movies. The findings offer important insights into the situation of audio description in China, especially in Shanghai, and recommendations are put forward for future developments to serve the community.

Keywords: accessibility studies; media accessibility; audio-visual translation, audio description; China; audio describers; user needs; user satisfaction; AD guidelines, text-to-speech; reception studies

Acknowledgements

I would like to start thanking the two persons that have accompanied me throughout the whole process: my PhD supervisors, Dr. Sara Rovira-Esteva and Dr. Helena Casas-Tost. I will never be able to express with words how grateful I feel to them for their trust and for their unconditional support. I would not have been able to finish this thesis without their encouragement and patience and I will always remember how easy they made my life during these (extremely challenging) years. I have learnt so much from them, not only academically but also from a more human point of view. They have inspired me and have taught me that being a good researcher is compatible with being kind, generous and humble.

I would also like to thank all the Chinese persons that have contributed to my research. From my one-semester Chinese tutor at Shanghai International Studies University, Dr. Xiao Weiqing, to all the audio description stakeholders that have assisted me. Without them, this would have never happened, and my research would not make any sense. Special thanks to my dear prof. Han Ying, prof. Qu Dapeng, prof. Jiang Hongyuan, prof. Huang Yiqing and the end users participating in the study. Thanks for trusting this *laowai* (foreigner).

My colleagues at MRA have also played a key role these years. Sharing our worries and achievements has been a source of great comfort during this journey, probably more than what they actually imagine. So, thank you, Belén, Dani, Gonzalo, Anita, Irene, Felipe, Joan and Blanca.

Last but not least, I would like to thank my family, especially my mother and my husband. Thank you for not losing your faith in me not even for one second, for your company, for your interest in my work, for providing doses of reality every time I lost hope and, in the case of Jose, for patiently waiting for me on the other side of the world.

To you all, thank you from the bottom of my heart.

Irene Tor-Carroggio

Barcelona, 22 April 2020

This research was supported by the European project EasyTV: Easing the access of Europeans with disabilities to converging media and content [GA: 761999], the RAD project (PGC2018-096566-B-I00) and the Confucius Institute thanks to a Joint Research Ph.D. Fellowship (2019) for this thesis' fieldwork.

Contents

Chapter 1: Introduction.....	15
1.1. Motivation and justification	16
1.2. Objectives and hypothesis	21
1.3. Theoretical framework.....	24
1.4. Methodology	31
1.5. Structure of the dissertation	37
Chapter 2. Article 1: Chinese audio describers' profile	41
2.1. Introduction.....	43
2.2. Methodology	46
2.3. Audio description in China	48
2.4. Profiling audio describers: Results of the questionnaires	54
2.5. Conclusions and proposals for the future	62
Chapter 3. Article 2: Chinese persons with sight loss' opinion on their experience with audio description	67
3.1. Introduction.....	68
3.2. Research background.....	71
3.3. Methodology	73
3.4. Results of survey 1.....	75
3.5. Results of survey 2.....	78
3.6. Discussion.....	82
3.7. Conclusions.....	85
Chapter 4. Article 3: Chinese AD guidelines	93
4.1. Introduction.....	95
4.2. European Guidelines and Standards	97
4.3. Methodology.....	100

4.4. Chinese AD Guidelines	101
4.5. Discussion.....	107
4.6. Conclusions and implications of the study	109
Chapter 5. Article 4: Reception study on the acceptance of TTS AD.....	115
5.1. Introduction.....	116
5.2. Previous research on TTS	118
5.3. Methodology	121
5.4. Results and discussion	125
5.5. Conclusion	135
Chapter 6. Discussion and conclusions	143
6.1. Discussion.....	144
6.2. Conclusions.....	153
6.3. Future lines of research.....	161
6.4. Limitations	164
6.5. Knowledge transfer to society	165
Chapter 7. Bibliography	169
Annexes	187
Annex 1: Articles within this dissertation.....	188
Annex 2: Additional articles	208
2.1. Additional article 1	208
2.2. Additional article 2	224
Annex 3: Ethics Committee Documentation	246
3.1. Procedure approved for the study with audio describers.....	246
3.2. Approval for the study with audio describers.....	248
3.3. Procedure approved for the study with users.....	249
3.4. Approval for the study with users.....	251
3.5. Procedure approved for the reception study	252

3.6. Approval for the reception study	254
3.7. Informed consent form for the study with audio describers (Chinese).....	254
3.8. Informed consent form for the study with audio describers (English)	255
3.9. Informed consent form for the study with users (Chinese)	255
3.10. Informed consent form for the study with users (English)	256
3.11. Informed consent form for the reception study (Chinese).....	257
3.12. Informed consent form for the reception study (English).....	259
Annex 4: Questionnaires.....	261
4.1. Questionnaire for the audio describers (Chinese).....	261
4.2. Questionnaire for the audio describers (English)	273
4.3. Questionnaire for the end users (Chinese).....	285
4.4. Questionnaire for the end users (English).....	289
4.5. Questionnaire distributed by Sound of Light (Chinese)	293
4.6. Questionnaire distributed by Sound of Light (English).....	295
4.7. Questionnaire for the reception study (A) (Chinese).....	298
4.8. Questionnaire for the reception study (A) (English)	301
4.9. Questionnaire for the reception study (B) (Chinese).....	303
4.10. Questionnaire for the reception study (B) (English).....	306
4.11. Post-questionnaire for the reception study (Chinese)	307
4.12. Post-questionnaire for the reception study (English).....	310
Annex 5: AD Scripts.....	313
5.1. Clip A's script.....	313
5.2. Clip B's script.....	314
Annex 6: Interviews.....	316
6.1. Interview 1 (Confucius Institute).....	316
6.2. Interview 2 (Shanghai International Studies University)	325

Index of Figures

Figure 1.1. Objectives of this research	23
Figure 1.2. Lewin's AR cycle (Source: Di Giovanni, 2018, p. 161).....	32
Figure 1.3. Need to investigate TTS AD	32
Figure 2.1. Fields in which audio describers have been trained.....	56
Figure 2.2. Audio describers' experience in hours	56
Figure 2.3. Frequency of collaborative work	57
Figure 2.4. Frequency of collaboration with persons with sight loss	57
Figure 2.5. Gateway to AD.....	59
Figure 2.6. Notion of AD	60
Figure 2.7. Perception of the qualities of being an audio describer	60
Figure 2.8. Audio describers' educational background.....	61
Figure 3.1. Informants' educational background	76
Figure 3.2. Fields in which respondents would like to have AD.....	77
Figure 3.3. Informants' preferred movie genres.....	78
Figure 3.4. Informants' education background	79
Figure 3.5. Informants' preferred movies.....	81
Figure 5.1. Experience with TTS.....	126
Figure 5.2. TTS frequency of use	126
Figure 5.3. Voice preferred.....	131
Figure 5.4. Interim acceptance	132
Figure 5.5. Permanent acceptance	133
Figure 6.1. Tasks performed by ADLAB PRO and our respondents	145
Figure 6.2. ADLAB PRO and our respondents' working experience.....	146
Figure 6.3. How audio describers work.....	150
Figure 6.4. Audio describers collaboration with users	151
Figure 6.5. How audio describers started working in this field.....	152

Index of Tables

Table 5.1. Clips selected and their characteristics	123
Table 5.2. Listening order of the clips	124
Table 5.3. Descriptive results	129
Table 5.4. Wilcoxon test results	130
Table 5.5. Comparison with similar studies	133
Table 6.1. Statements regarding AD	148

Abbreviations and Acronyms Glossary

AD	Audio description
AI	Audio introduction
AR	Action research
AS	Accessibility Studies
AST	Audio subtitling
AVT	Audio-visual Translation
CRPD	Convention on the Rights of Persons with Disabilities
EU	European Union
H	Research hypothesis
ISO	International Organization for Standardization
ITU	International Telecommunications Union
MA	Media accessibility
MOS	Mean opinion score
SDH	Subtitles for the deaf and hard of hearing
SLI	Sign language interpreting
TTS	Text-to-speech
TS	Translation Studies
UAB	Universitat Autònoma de Barcelona
UK	United Kingdom
UN	United Nations

Chapter 1: Introduction

Many two-footed people laugh at me for having one foot, which always used to infuriate me. But as soon as I arrived here at our master's place, my rage fell away and I returned to normal. It's as if the master had cleansed me with his goodness without my even realizing it. I have studied under him for nineteen years and never once have I been aware that I was one-footed. Here you and I wander together beyond shapes and bodies, is it not wrong of you to seek me within a particular body and shape?

Zhuangzi: Essential Writings with Selections from Traditional Commentaries (5.12)

(Translated by Brook Ziporyn)

1.1. Motivation and justification

In March 2019, the Chinese film director Jia Zhangke (贾樟柯) submitted a motion in the 13th National People's Congress to advocate the implementation of what in China¹ is known as *wuzhang 'ai dianying*² (无障碍电影, barrier-free movies). Jia submitted a proposal concerning the need to legally solve the copyright challenges that “barrier-free movies” face. He also underlined the importance of regulating and standardising media accessibility (MA). That was the first time a film director has publicly defended the existence of sensory access services to grant people with disabilities the right to enjoy movies.

The perception of disability has come a long way in the last decade in China. Events such as the Beijing Summer Paralympics and the Sichuan earthquake, both in 2008, have contributed to changing how the Chinese society views people with disabilities, which has usually been influenced by superstitions. In fact, there are numerous examples that illustrate how little value this group of people has traditionally held, for instance, the Chinese one-child policy giving couples a second chance if their first child was born with a severe disability. Nowadays, and even though the situation has considerably improved, the life of Chinese people with disabilities is still far from ideal. Not only do prejudices and superstitions persist, but the huge domestic interregional socioeconomic differences also have an especially severe impact on this extremely diverse group of people. This is

¹ This thesis will refer to the People's Republic of China as “China”. In cases where regions in Greater China are mentioned, such as Hong Kong and Taiwan, we refer to them individually.

² This term reminds the author of that of “barrier-free design”, a concept that appeared in the 1950s to address the issue of access in a built environment for persons with disabilities. This concept evolved and it has now been substituted for “universal design”, which focuses on providing access to all individuals.

due to most people with disabilities in China living in rural areas, which means that they must deal with multiple disadvantages, such as geographical isolation, lack of adequate infrastructure and services, and lower incomes than those in urban areas. The latest official statistical data regarding disability, which unfortunately dates back to more than ten years ago, evidences the problem's magnitude. According to the 2006 Second China National Sample Survey on Disability (China Disabled Persons' Federation, 2007), at least 82.99 million people suffer from one or more disabilities in China, accounting for 6.34% of the total population. There are slightly more men with disabilities (51.55%) than women (48.45%) and just a quarter of all these people live in cities (24.96%). This stands in stark contrast with, for example, the Spanish case, in which only a quarter of people with disabilities live in the countryside (i.e., places in which a maximum of ten thousand people live) (Observatorio estatal de la discapacidad, 2018). This data, however, needs to be contextualised: while only 16.5 % of Spaniards reside in cities (Efe Agro, 2018), 40.42% of Chinese citizens still live in rural areas (Xinhua News Agency, 2019). Going back to the official data from 2006 (cited in Campbell & Uren, 2011), over 50% of Chinese people with disabilities are 60 years old or above (53.24%) and 42.10% are between 15 and 59 years old, whilst the percentage is reduced to 4.66% in the 0-14 years age-group. With regards to their income, urban households with a disabled family member have a per capita income of 4,864 yuan. The number goes down to 2,260 yuan in rural areas. It is also important to mention that for 12.95% of rural households with a disabled member, their income was below 683 yuan and that for 7.96% of rural households with a disabled member their income was between 684 and 944 yuan. In fact, disability is linked to poverty, since, of the 30 million Chinese living in poverty, 80% are disabled.

The socioeconomic profile of this group can explain why the Chinese government has prioritised meeting certain needs of people with disabilities at the expense of others. Nonetheless, since China ratified the United Nations' Convention on the Rights of Persons with Disabilities (UN CRPD) in 2008, some existing laws have been amended to foster respect for the rights and dignity of people with disabilities and to allow them to participate fully in some aspects of life. For example, China has worked on the legal protection of workers with disabilities from exploitation, violence and abuse, with their efforts captured in the Law on the Protection of Persons with Disabilities, the Law on Public Security Administration Punishments and the Law on Employment Contracts. For these improvements, among others, China was congratulated in the first and only review

performed by the UN in 2012. Notwithstanding, there are still many challenges that China, as the country with the highest population of people with disabilities in the world, has to address, such as the provision of sensory access services in media content.

In China the media industry is not just an arena for traditional media to compete; other players include mobile phone media, digital magazines, mobile TV, the Internet and digital TV (Price Waterhouse Coopers, 2016, p. 8). With the emergence of a new culture, entertainment and media industry, China's policies have become more regulated, but sensory access services are not even considered as an afterthought. Notwithstanding, and interestingly enough, Chinese people with disabilities believe there are other more pressing issues to tackle, such as employment (Li, 2013). Yet, in the digital and information era, media is already indispensable if you want to participate in society, no matter how old you are and whether you live in a city or in the countryside. Price Waterhouse Coopers (2016) conducted a study in which entertainment behaviour was analysed in China. The study revealed that Chinese youths are dependent on mobile phones and are "more involved in passive engagement with information that is pushed on them rather than reading to proactively obtain information" (Price Waterhouse Coopers, 2016, p. 3). This means that TV and movies are excellent tools for reaching them. Young people's affinity with social media is also an important characteristic to bear in mind due to the predominance of audio-visual content in it. The case of *Taizifei Shengzhi Ji* (太子妃升职记, Go princess go), a low-budget web series launched on social media platforms, illustrates the potential of social media achieving 200 million clicks in a single day. In this era of mobile internet applications, "users are expecting to enjoy communications at anytime, anywhere" (Price Waterhouse Coopers, 2016). In addition, adapting online novels into TV series and movies has become increasingly popular. More specifically, by the end of 2014, 114 online novels had been adapted in this way (Price Waterhouse Coopers, 2016, p. 5). In addition, "China's box office revenue and number of movie-goers are expected to surpass North America by 2020" (Deloitte, n.d., p. 2). Although this information is mainly related to the young, it cannot be forgotten that they are the future users of sensory access services as disability is closely related to age (Peng *et al.*, 2010).

Sensory access services are not completely unknown in China. These services are targeted, although not exclusively, at people with sight or hearing impairment and include subtitles for the deaf and hard of hearing (SDH), sign language interpreting (SLI) and audio description (AD). One of the main areas in which they can be applied is in TV content.

Yet, Chinese people with sensory disabilities do not have proper access to this TV content, since Chinese TV lacks –or does not offer enough of– basic access services that are crucial for it to be accessible on an equal basis. Although subtitles are required for almost all TV programmes by law, these do not specifically cater for the needs of people with hearing loss but cater for other needs (Casas-Tost & Rovira-Esteva, 2018). Regarding SLI, it has become a relatively widespread type of media interpreting in China, with more than 170 TV channels broadcasting sign-language-interpreted programmes daily or weekly (Xiao & Li, 2011). Yet, Xiao and Li's (2011) study evidenced the current problems which SLI on Chinese TV faces, such as low comprehension rates, interpreters being selected by virtue of their looks, and SL interpreted programmes being scarce and broadcast at rather inconvenient airtimes. With regards to AD, it is completely non-existent on TV, although it is available in other audio-visual products, mainly movies.

AD, which is what we will be placing most emphasis on in this thesis, is a “description of visual information delivered via an audio channel” (Salway, 2007, p. 151); it can also be defined as “the addition of a descriptive narrative to accompany the key visual elements of theatre, television, cinema and other visual media” (Greening & Rolph, 2007, p. 127); or simply “the translation of images into words” (Bourne & Jiménez-Hurtado, 2007, p. 176). AD is currently mainly offered in the so-called “first-tier cities”, such as Beijing, Shanghai and Guangzhou. This is not surprising considering that the provision of AD relies on the priorities of each regional government, along with grassroot initiatives (Ma, 2016). There are 17 million people who suffer from sight loss in China, but this access service is only available to 5.8%-8.8% of them (Ma, 2016, p. 111). In fact, Ma (2016, p. 111) complained about the number of audio described movies not even reaching 100 in 2016. This meant that people with sight loss were unable to access more than 90% of the movies released in China. Despite AD's paucity, Ma (2019, p. 255) underlined the importance of granting people with sight loss access to movies, highlighting the various advantages that watching a film has to offer. For example, it is a cultural activity and a pastime. It is also a social activity thanks to which people with sight loss can gather and alleviate their sense of loneliness. Moreover, it is a way of discovering society and all the elements it comprises. In addition, the living conditions of people with disabilities have a direct impact on the happiness of around one-fifth of Chinese families and their 200 million relatives (Campbell & Uren, 2011). Yet, despite its importance for people with sight loss (and for their relatives), the practice of AD is lagging behind that of Taiwan

and Hong Kong. Nonetheless, these two regions have also been overtaken by other countries such as the USA and many European countries (Li, 2019, p. 107).

The academic side of AD does not come off better, although the state of Audio-visual Translation (AVT) is quite vibrant in China (Gambier & Jin, 2018). This is partly thanks to the current AVT policies in the country, which encourage both the translation of foreign films into Chinese and the translation of outstanding Chinese films into other languages, including Chinese ethnic minority languages. Yet, full-time audio-visual translators are a rarity and they are mainly employed by state-owned AVT studios. Furthermore, the emergence of online video platforms such as Iqiyi and Youku Chinese has resulted in a higher demand of AVT services, although fansubbing, i.e., amateur translation, is also resorted to in these cases (Gambier & Jin, 2018). Chinese AVT's most popular modalities (dubbing and subtitling) are mainly linked to interlingual translation leaving little room for MA services.

In spite of AVT education being offered in China in universities such as Shanghai International Studies University and China Communication University and the number of academic articles on AVT increasing dramatically since the 1990s (Gambier & Jin, 2018), the academic study of AD in China, especially on the Mainland, is almost non-existent.³ This is probably due to Chinese scholars' disability-related research primarily focusing on basic needs such as the welfare system and education (Dauncey, 2007), since, as it has already been pointed out:

At present, there are still more than 10 million rural disabled persons without adequate food and clothing, and more than 20 million urban disabled persons with income below the minimum standard of living... and there is a big gap between their living conditions and the average level of social living conditions. (Jia & Zhao, 2008, p. 3)

With the exception of Li (2013), Chinese academia is still oblivious –or turning a blind eye– to one of the hottest research fields in AVT. Yan and Luo (2019, p. 10) observe that the few Chinese scholars tackling AD highlight the socialist values that it embodies. They also suggest how to widen its availability or stress the differences in terms of academic research that exist in comparison with the West. This stands in sharp contrast with other TS fields, which have made great progress in the last few years, such as machine translation and automation technologies (Gambier & Jin, 2018). This weakness has

³ Yet, there are some universities such as Qinghua University and Tianjin University that already offer courses on accessibility from a more architectonic and engineering point of view.

already been admitted by some scholars, such as Gambier and Jin (2018) and Xiao and Peng (2019). It is also worth mentioning that the few articles and books on AD published by Chinese scholars that we have found are all written in Chinese, such as Li (2013) and more recently Yan and Luo (2019). This makes the academic debate between the Chinese and international community more difficult and is probably why western MA researchers have been unaware of the Chinese initiatives launched so far.

In the light of the precarious state of AD and the challenges it still has to face in China, we propose a thesis that not only contextualises AD in the country as a first step to launch academic research on the topic, but also offers an AD-related alternative that will facilitate the provision of AD while the country raises awareness of the importance of MA services. We believe this thesis is necessary to blow away the academic cobwebs that cover this field to encourage Chinese academia to engage in this type of research, which most probably could result in interesting and reliable data to inform legal and governmental decisions that can translate into improving the lives of Chinese people with disabilities.

Nonetheless, and since “the attempt to present an overall picture for anything at all about China often provokes trepidation before any other intellectual responses” (Yeung, 2007, p. 231), we must warn the reader to bear in mind that the consequences of the uneven economic development across China hinders our ability to making general assumptions about the country. Therefore, our results cannot be divorced from their geographical source, which are usually first-tier cities.

1.2. Objectives and hypothesis

This section describes the objectives that motivate our thesis and the hypothesis we departed from.

The first general aim of this dissertation is:

1. To describe the history and the current state of AD in Mainland China, both as an access service and as an object of study, with special focus on the city of Shanghai.

In order to achieve this objective, the following specific objectives are defined:

- a) To present the history and the current state of AD in China, both as an access service and as an academic research field, including the legal framework within which this access service is framed.
- b) To profile audio describers in Mainland China.
- c) To investigate the habits, needs and satisfaction of AD users towards this access service.
- d) To analyse some Chinese AD guidelines currently in use.

The second general aim of this dissertation is:

- 2. To compare the Chinese case with the European context in some specific aspects.

In order to attain this objective, the following specific objectives are established:

- a) To compare the Chinese audio describers' profile with that of the Europeans'.
- b) To compare the Chinese AD users' habits, needs and satisfaction with the European AD users'.
- c) To compare the Chinese guidelines under study with the European trends in this regard.

The third general aim of this dissertation is:

- 3. To determine whether text-to-speech (TTS) AD can be accepted in movies in China.

In order to accomplish this objective, the following specific objectives are formulated:

- a) To evaluate TTS AD in Chinese by comparing it with standard human-voiced AD on key features.
- b) To elucidate whether TTS AD could be accepted as either an interim or a permanent solution, or both, in the Chinese context.

The objectives are summarised in Figure 1.1:

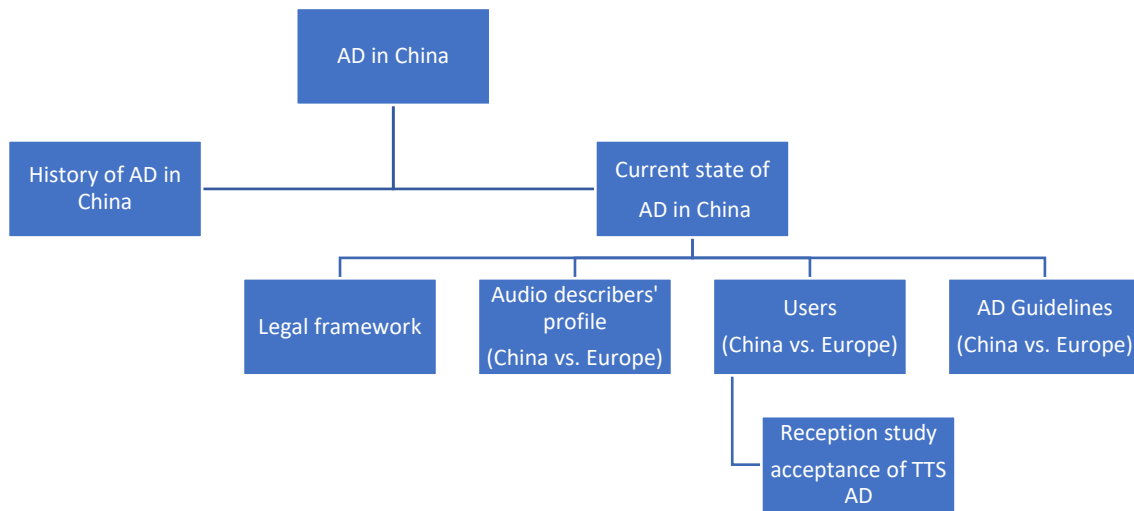


Figure 1.1. Objectives of this research

As it can be seen, we aim at delving into the past (history), present (current state of the service) and future (alternatives to speed the production of AD) of this access service in China. Related to these aims, the following research hypothesis (H) have been formulated:

- a) H₁: AD is an access service in its infancy in China and an under-researched topic in Chinese academia.
- b) H₂: Audio describers are more experienced and professional, users more demanding and guidelines more exhaustive in Europe than in China.
- c) H₃: Chinese AD users will prefer human-voiced AD but will accept TTS AD both as an interim and permanent alternative.

Each of these aims is tackled in a different paper. The first study, entitled “Who is currently audio describing in China? A study of Chinese audio describer profiles” addresses the first general aim in Chapter 2. This article, along with the one in Chapter 3 and the one in Chapter 4, entitled “The customer is always right: Study on Chinese people with sight loss’ opinion on their experience with audio description” and “When East meets West: A comparison of audio description guidelines in China and Europe”, respectively, contextualise AD in China and establish a comparison with the situation in Europe, tackling the second aim. Finally, the fourth paper, entitled “T(ime) T(o) S(tart) synthesising AD in China? Results of a reception study”, concentrates on the reception

study carried out in Shanghai. It is presented in Chapter 5 and it tackles the third general aim of this dissertation.

1.3. Theoretical framework

“Having clarity around the perspective from which one conducts her analysis is critical to avoid being stuck in a conceptual, epistemological, or methodological muddle” (Greco, 2019, p. 22). Given the fact that the ultimate goal of this thesis is to provide more access to AD to Chinese people, this PhD research is, at its core, embedded in the field of Accessibility Studies (AS), which investigates, among others, accessibility processes and phenomena (Greco, 2019).

The independent existence of AS is still questioned by some researchers, but Greco (2019) claims it already gathers all the elements required for a new field to emerge, i.e., interdisciplinarity, the formation of a research community and forms of opposition of well-established fields. In effect, Greco (2019, p. 22) argues that AS have been, “de facto, a field for some time”. He explains that it is common for different disciplines to collaborate and “split apart unchanged when the work is done” (Greco, 2019, p. 20), but in the case of MA the issues under study require the integration, not just the juxtaposition of different fields. This integration has resulted in MA experts having more in common among themselves than with the colleagues from the fields they originally come from (Translation Studies, Psychology, Performing Studies, Engineering, Computer Science, to name a few), but their emancipation still faces opposition. A concrete example: the author of this thesis, who has a background in Translation has more in common with the members of the Enhancing Audio Description Project,⁴ who are mostly electronic engineers and art and music experts, than with researchers investigating machine translation.

AS is also the umbrella under which MA can be found. According to this approach, MA focuses on “concerning access to media and non-media objects, services and environments through media solutions, for any person who cannot or would not be able to, either partially or completely, access them in their original form” (Greco, 2019, p. 18).

⁴ This project (<http://enhancingaudiodescription.com/>, last accessed 5 March 2020) aims at reducing the number of AD verbal descriptions by using innovative sound techniques.

Despite ultimately belonging to AS, Greco (2019, p. 19) defends that this universalist account of MA:

[...] favours a convergence of the different conceptions of MA and of MA services developed in other fields. In turn, this positions MA as a broader interdisciplinary area that criss-crosses many fields, including AVT, but that cannot be entirely nor exclusively reduced to any of them”.

Therefore, it can be affirmed that AVT (and, therefore, TS) overlaps in some aspects with MA. Yet, Greco (2019) admits that speaking of “overlapping” is a simplistic way of saying that some problems can both be analysed from the perspective of TS and AS. That said, “not all translation problems are accessibility problems; and not all accessibility problems are translation problems” (Greco, 2019, p. 23).

Despite the evidence proving the current existence of AS, AVT is the field where MA has been –and still is in many cases– developing as a research discipline. Traditionally, MA was considered to be a subarea within AVT dealing with access services targeted at people with sensory disabilities, and has recently become a synonym for AVT (Greco, 2019). This equivalence originates from the inherent –and more recent– purpose of AVT being to grant access to audio-visual content no matter what personal characteristics individuals have. This means that, for example, a person that cannot speak Hungarian will obviously need an AVT modality to access a Hungarian movie as much as a Hungarian person with hearing loss will need SDH for the very same movie. In other words, this person not speaking Hungarian is as “disabled” as a person with hearing loss because neither can access this movie unless an AVT modality is provided. Moreover, if the person reading the subtitles is a slow reader, he or she may also be regarded as “disabled” because, compared to faster or average readers, regular subtitles will still not grant him or her access to that specific audio-visual content. Thus, disability can also be conceptualised as a melange of factors that do not necessarily involve physical disability (Mitra, 2006). In the same vein, Gambier (2018) unties accessibility from disability by pointing out that “[a]ccessibility is also a key concept that shakes up the dominant way of assessing quality of a translation and it covers many aspects” (p. 55), such as acceptability, legibility, readability, relevance, among others. This more holistic approach has been welcomed by academia (see for example, Gerber-Morón, 2018; Romero-Fresco, 2018; Agulló *et al.*, 2018) and the study of the so-called “classic AVT modes” (namely dubbing, subtitling and voice-over) now shares predominance with that of sensory access services, such as SDH, AD, SLI, audio introductions (AI), audio subtitling (AST) and

clean audio, to name the most relevant ones. This point can also be illustrated by doing a quick search in the Media Accessibility Platform,⁵ a bibliographical database aiming at collecting, among others, all academic publications devoted to AVT, including MA. At the time this thesis was conducted, there were 324 entries for subtitling-related publications, 289 for publications dealing with SDH and 535 for publications revolving around AD.

Despite us initially admitting that our research is framed within AS, the author of this dissertation has decided to implement her research through TS, which is the field in which, as we have already pointed out, MA has traditionally been developing in. This is due to the inclusion of MA within AS being in its infancy in terms of academic literature and given the latest and above-mentioned paradigm shift in AVT. In addition, Greco and Jankowska (forthcoming) include AD among the “translation-based” access services, unlike, for instance, clean audio.

In this dissertation we define AD, as a “semiotic rendering of images into words” (Hernández-Bartolomé & Mendiluce-Cabrera, 2009, p. 5), dealing with audio-visual texts, in which “various components combine to generate one same meaning” (Hernández-Bartolomé & Mendiluce-Cabrera, 2009, p. 4). These components include different channels (the acoustic and the visual) and various codes (dialogues, costumes and make-up, moral attitudes, proxemic and kinesic elements, and cinematographic techniques, just to name a few). Despite the initial doubts of whether to underpin MA services under the umbrella of TS since they do not necessarily involve interlingual transfer (Díaz-Cintas, 2005, p. 4; Gambier & Jin, 2019, p. 222), the belief that AD is a type of intersemiotic translation has become commonplace. The fact that it shares some constraints (e.g. time/space constraints) with other AVT modalities has also been decisive in coming to this conclusion, although Braun (2008, p. 11) claims that “live AD may be better conceptualised as audiovisual interpreting rather than translation”. In any case, AD has also been tagged as both an intermodal (Braun, 2008) and a multisemiotic (Szarkowska & Orero, 2014) type of translation. We shall now briefly explain why AD fits in all three categories (intersemiotic, intermodal and multisemiotic translation).

First, AD is considered to be a type of intersemiotic translation because it entails a transfer between modes: from the nonverbal (images, sound, etc.) to the verbal. The term

⁵ See <https://mapaccess.uab.cat/> [Last accessed 3 October 2019]

“intersemiotic translation” was coined by Jakobson (2000 [1959]) and, although the original definition “assumes that the interpretation of verbal signs is done by means of signs of nonverbal sign systems” (Walczak, 2017b, p. 23), AD can still be made to fit in that category, albeit when the semiotic transmutation is performed in reverse order. Jakobson’s (2000 [1959]) classification of types of translation also included intralingual and interlingual translation, categories under which AD can also fall (Hernández-Bartolomé & Mendiluce-Cabrera, 2009). For example, AD can involve interlingual translation in dubbing and subtitling countries, since this access service requires “an interlinguistic translation prior to AD itself” (Hernández-Bartolomé & Mendiluce-Cabrera, 2009, p. 5). When preparing AD, it may also be necessary to add extra explanations to make a piece of nonverbal information clearer. This may be important in many cases given that “non-verbal communication is essential in the understanding of a film. Therefore, it should be audio described” (Igareda, 2011, p. 237). It is precisely in these cases in which intralinguistic translation may also come into play. Second, AD can be viewed as multimodal translation for three main reasons: its creation involving a cross-modal transfer, the source text being multimodal and AD interacting with other codes, such as music or sound effects (Braun, 2008). Last but not least, AD can even be regarded as a multisemiotic type of translation for the number of channels involved in the transfer of information. Szarkowska and Orero (2014, p. 124), using Gottlieb’s (1998, in Szarkowska & Orero, 2014) classification of communication channels, claim that AD switches from more than one channel (such as the verbal visual channel in the case of on-screen text or the nonverbal visual channel in the case of costumes) to the verbal auditory channel. As it can be seen, the feature that stands out the most in AD, and that all terms reflect in one way or another, is the transfer between the visual and the verbal.

Although not exclusively, AD is still mainly targeted at people with sight loss (ISO, 2015). This has some obvious implications, the main one being that its special uses need to be taken into account:

On the one hand, AD aims at social integration, particularly of visually impaired people, but also of other social groups; on the other hand, audio described material may help people broaden their minds, as describers apprise viewers or listeners of some cultural references certain programmes sometimes contain. In any case, the enjoyable element is crucial. (Hernández-Bartolomé & Mendiluce-Cabrera, 2009, p. 3)

These three dimensions (inclusion, cognition and entertainment) of AD demand a better knowledge of viewers' needs, habits and reception capacity (Gambier, 2018). This can be achieved through reception studies, which are the current focus of numerous MA researchers (Walczak, 2017b; Leung, 2018; Gerber-Morón, 2018; Iturregui-Gallardo, 2019). Díaz-Cintas and Szarkowska (2020) also acknowledge the “unprecedented boom in experimental research conducted in the field of AVT” to gain a deeper understanding of the audience's behaviour and attitudes towards the consumption of translated audio-visual productions. Greco (2019, p. 19) underlines one of the main reasons why reception studies have gained more importance: “users are increasingly seen as bearers of valuable knowledge for the investigation of accessibility processes and phenomena”. This, although might seem new, was already brought up by Plato:

It necessarily follows, then, that the user of anything is the one who knows most of it by experience, and that he reports to the maker the good or bad effects in the use of the thing he uses. [...] Then in respect of the same implement the maker will have right belief about its excellence and defects from association with the man who knows and being compelled to listen to him, but the user will have knowledge. (*Republic* 601d1–602a1)

The special issue of the *Journal of Specialised Translation* published in 2020 and focusing largely on AVT is a clear illustration of this momentum. This stands in stark contrast with the early studies on AD, which were descriptive (Benecke, 2004; Snyder, 2007; Bourne & Jiménez-Hurtado, 2007; Palomo López, 2010; Cabeza-Cáceres, 2010; Rodríguez Posadas, 2010). According to Díaz-Cintas and Szarkowska (2020, p. 4), “this evolution is a natural one that mimics similar developments in the wider field of translation studies [...]. Yet, the difference [...] resides in the object of study”. While in TS the focal point has traditionally been the translation process and the behaviour of the translators themselves, AVT, and therefore, MA, has focused more on the receptors (i.e., the viewers) and the practitioners (Díaz-Cintas & Szarkowska, 2020).

Regarding reception studies, the concept of “reception” attracted attention in TS after Nida (2000, in Gambier, 2018) underlined the pivotal role of readers in interlingual communication. Since then, reception has played a key role in TS, although such an important dimension still merits more research (Gambier, 2018). The importance of studying the reception of translation lies in the key insights obtained that enable translators to produce a better-quality output. In this case, “quality” is “determined by whether readers respond to it in the same way in which readers responded to the ST”

(Gambier, 2018, n.p.). Reception has also been approached from different fields, such as Sociology, Applied Linguistics and Social Psychology and it is precisely this multiplicity of perspectives that hampers the detection of a reception subfield in TS (Gambier, 2018). Gambier (2018) specifies what is understood as “reception studies” in the AVT context:

Studying reception means to investigate the way(s) in which AV products/ performances are processed, consumed, absorbed, accepted, appreciated, interpreted, understood and remembered by the viewers, under specific contextual/socio-cultural conditions and with their memories of their experience as cinema going. (Gambier, 2018, p. 56)

In the broad context of TS, reception studies can be approached from three different angles, namely, response, reaction and repercussion (Chesterman, 2007, p. 70), which Gambier (2018, p. 55) calls “the 3 Rs”. Studies focusing on the response are those revolving around the perceptual decoding of a translation. This is investigated mostly by experimental psychologists (Gambier, 2018, p. 57). Studying “reaction” means observing the mental processing and understanding of a translation (Tuominen, 2018, p. 75). Studies focusing on the reception of humour or cultural references are wonderful examples (see, for instance, Iaia, 2015 or Chiaro, 2007). With regards to “repercussion”, this concept is understood as “viewers’ attitudes as well as sociocultural consequences” (Tuominen, 2018, p. 70). These “attitudes” are explored more globally and “they provide context for studies more oriented towards (...) an understanding of the role of audiovisual translations in society, media and culture” (Tuominen, 2018, p. 77). The work carried out by Szarkowska (2011) and Jankowska (2015), for instance, are examples of reception studies focusing on the repercussions on the objects under study. Our thesis is also under the “repercussion” umbrella, since it intends to ascertain Chinese AD users’ attitude towards TTS AD and the sociocultural consequences that can follow the introduction of such a service.

The reception of the most traditional AVT modes has been widely researched, especially from the “reaction” and “repercussion” perspectives. To name just a few, see Perego *et al.* (2010), Lavaour & Bairstow (2011), Orrego-Carmona (2015) and Nikolic (2018) for subtitling; Koverienė & Satkauskaitė (2018) and Di Giovanni & Romero-Fresco (2019) for dubbing; and Di Giovanni (2012) and Matamala & Ortiz-Boix (2018) for voice-over. The studies revolving around the reception of sensory access services are also increasingly numerous, including, for example, Wehrmeyer (2014); Romero-Fresco (2018) and Iturregui-Gallardo (2019). In our case, the works studying the reception of

AD are the most relevant, since AD is the core of our thesis. The reception of AD has been studied when applied to, for example, opera (Eardley-Weaver, 2014) and fashion shows (Udo & Fels, 2009). Yet, filmic AD is the most widespread and researched AD type (Perego, 2019, p. 112). Consequently, the studies revolving around its reception are the most numerous. The parameters or aspects studied are wide-ranging and include, without purporting to be exhaustive, the requirements of people with sight loss in order to develop a system to make television accessible through AD (Lodge, 1993); the advantages of film AD for people with sight loss (Peli *et al.*, 1996; Schmeidler & Kirchner, 2001); the differences between ADs for the same film in different languages (Bourne & Jiménez-Hurtado, 2007; Matamala & Rami, 2009); the potential demand for audio described Bollywood films in the UK and India (Rai, 2009); the reception of films in which AI and AD (Romero-Fresco & Fryer, 2013) have been combined; the paralinguistic aural stimuli in the assessment of AD (Iglesias Fernández *et al.*, 2011); the effect of speech rate on user comprehension (Cabeza-Cáceres, 2013); the emotional impact of language in AD (Ramos, 2013; 2015); the amount of information favouring AD users' memory regarding character description (Fresno, 2014); various AD strategies for specific issues to detect preferred solutions (Chmiel & Mazur, 2016; Leung, 2018); cheaper and faster ways to produce AD such as TTS AD (Fernández-Torné, 2016) and AD translation (Jankowska & Fryer, 2017); the impact of AD on the emotional response of AD users (Walczak, 2017b); the reception of foreign words (Sanz-Moreno, 2018); the application of AD in 360° videos (Fidyka, 2018); the integration of AD in the creation process of audio-visual content (Fryer, 2108); and the cultural competence of the audience in AVT, particularly in AD (Sanz-Moreno, 2019).

Di Giovanni (2018b) offers a more systematic classification of reception studies on AD, which she categorises into four different strands, namely:

- a) The what-to-describe strand: studies investigating what the object of descriptions should be. A substantial number of AD-related reception studies belong to this category, such as those derived from The Pear Tree Project (Orero, 2008).
- b) The psychology-based strand: research based on psychological methodologies, such as that of Iturregui-Gallardo (2019), who studied how voicing strategies affected the activation of emotions. Other examples are the studies conducted by Fryer and Freeman (2014), who focused on presence and measured emotional experience and Fryer and Walczak (2017), who investigated the emotional

reaction to emotion-conveying AD. Di Giovanni (2018b, p. 241) notes that this strand is extremely influential.

- c) The alternative routes strand: studies aiming at producing AD using unconventional methods such as AD translation (Jankowska, 2015) and TTS AD (Fernández-Torné, 2016).
- d) The inclusive strand: a research line viewing AD as part of universal design in, for example, museums (Neves, 2016).

This variety of strands and the multiplicity of studies dealing with different aspects of AD evidence that AD has already come of age.

1.4. Methodology

The methodology followed in each stage of our research has been clearly explained in the corresponding articles. Therefore, in order to avoid unnecessary repetition, the emphasis here will be put on other issues that have not been mentioned or were succinctly described in the articles. This is the case of the general methodological approach that underlies in this thesis, namely the action research (AR) cycle.

AR is a form of research “intended to have both action and research outcomes” (Dick, 1993, n.p.). Its “action” side aims at bringing about change in some community, whilst its “research” facet seeks to broaden understanding “on the part of the researcher or the client, or both” (Dick, 1993, n.p.). Despite AR having a twofold objective, some of the methods this approach encompasses tend to emphasize one facet or the other. When the “action” side is stressed, research remains as a “fringe benefit”, while in the cases where research is the main focus, publication to reach a wider audience of researchers is sought. In the latter, “more attention is often given to the design of the research than to other aspects” (Dick, 1993, n.p.). Yet, no matter what approach is highlighted, in both cases it is possible for “action to inform understanding and understanding to assist action”. Additionally, this form of research always takes the shape of a spiral, in which planning precedes action and review follows (Figure 1.2.).

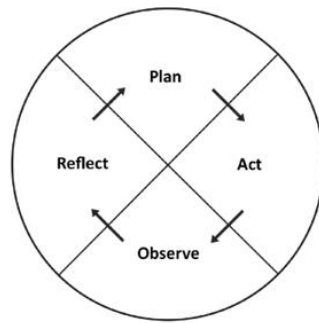


Figure 1.2. Lewin's AR cycle (Source: Di Giovanni, 2018, p. 161)

This cycle seems to suggest that AR, and the participatory activities it is concerned with, are never an end in themselves, but rather the motivating force for further planning, acting, observing and reflecting (Di Giovanni, 2018). This cycle was originally conceived as a means to solve social problems by participative intervention, as well as to “raise the self-esteem of minority groups to help them seek independence, equality and cooperation” (Adelman, 1993). Di Giovanni (2018a) translates Lewin's “democratic participation” into “inclusion” and, given that our main aim is to give some food for thought to the Chinese MA stakeholders, this model allows us to complete it once and pass it on to those who will start the planning all over again.

Furthermore, our reception study followed a user-centric design, in which users decided what materials would be tested. They were also asked how necessary they thought investigating TTS AD was. Users were also given a questionnaire aiming at getting to know their habits and satisfaction with AD (see Chapter 3). This piece of information could not be added in the resulting article for the sake of brevity but is now presented in Figure 1.3.

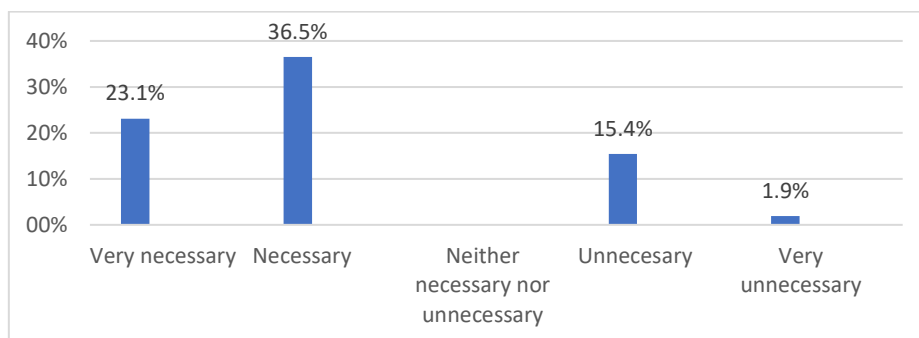


Figure 1.3. Need to investigate TTS AD

As it can be seen, more than half of the users ($N = 52$) participating in our study agreed with our idea of making AD more available in China through TTS being a worth researching topic. It was assumed that those not finding TTS AD a topic in need of investigating had other priorities, which could be explored in future studies. Yet, and although asking users' opinion is always necessary, Iglesias Fernández (2010, p. 216) believes that “reception studies in AD should not stop at users' preconceived preferences but further expose them to actual AD products for assessment”.

Implicating users was seen as a way to respond, at least in the European case, to a social call to involve people with disabilities in everything that concerns them. This is in line with the UN motto “nothing about us without us”. An example of the focal role users play can be seen in the current ISO AD standard (“TS 20071-21 ‘Information Technology — User interface component accessibility — Part 21: Guidance on audio description’”), which encourages the inclusion of consumers in the process of creating AD and during its revision. The latest EU-funded MA-related projects take the user-centric approach as well.⁶ Adopting this approach obviously has further implications, such as our dependence on users –which adds difficulty to our thesis due to the specific profile of the users needed and to this type of research being a novelty in the country– and it requiring that all procedures, from user selection to results dissemination, comply with ethical requirements. This has been successfully achieved in this thesis.

We would like to highlight the difficulty of recruiting Chinese users for a foreigner, especially taking into account that Chinese people with sight loss do not usually leave their homes. Most of the information gathered is not available on the Internet and various long stays in China were needed to find users and associations and build a relationship of trust with them. Despite the difficulties encountered, the good will with which we were welcomed and the collaboration they offered must be acknowledged and has been very much appreciated.

There is one point which has already been mentioned in all the publications included in this dissertation, but we would like to expand on it further: the ethics-related requirements when carrying out tests with users with disabilities. As Orero *et al.* (2018) observe:

⁶ See, for instance, the EasyTV Project (<https://easytvproject.eu/>, last accessed 26 April 2020) and the ImAc project (<https://www.imac-project.eu/>, last accessed 26 April 2020).

The growing interest in reception studies, the widespread use of both behavioural and physiological measures, the growing interest in media accessibility and the consequent involvement of vulnerable audiences in AVT experimentation (e.g. deaf and hard of hearing, blind and visually impaired, elderly, and children) highlight the need for norms for conduct that can guide researchers in actual research situations. Furthermore, applying for ethical approval is becoming an increasingly common step in the execution of AVT research projects. (Pérez-González, 2014 in Orero *et al.*, 2018, p. 110)

Before conducting any AVT study, researchers need to be given the green light from the ethics committee of their university. Orero *et al.* (2018, p.110) explain the aspects this typically implies:

Participants should give informed consent; data should be anonymised; data should be stored in a secure place for a set period (typically 5 years); privacy, perceived and real benefits from a study, and other relevant considerations should be considered and reported. Publications should report the status of ethics applications and clearance.

When applying for the ethics approval for the reception study, the Ethics Committee of the Universitat Autònoma de Barcelona (UAB) asked us to request that the Chinese university we were staying at at the time (Shanghai International Studies University) approve the reception study we had designed. This was not possible due to the university not having a committee for these purposes. Finally, our one-semester Chinese tutor (Dr. Xiao Weiqing) had to take the responsibility for it and declare so in a letter that was sent back to UAB's Ethics Committee. This situation brought to light once more that Chinese AVT research is still behind other countries and some practices and procedures need to be revised and systematised to ensure participants' rights are respected as well as to guarantee their safety.

Regarding the part of our thesis profiling audio describers and investigating user needs and satisfaction, no detailed explanation is included in any of the related publications as to why the ADLAB PRO questionnaire was deemed to be appropriate as a research tool in our thesis. The reason behind our choice lies in us aiming at helping expand the scope of the ADLAB PRO project by providing results from a non-European country that was not included in their study, which we hoped would allow for a better understanding of the current position of AD in the world. Also, since their “[...] secondary objective [...] is to gain a better picture of the sociological reality audio describers live in” we thought this was in line with what we also aimed at knowledge on regarding audio describers, i.e., their status, working experience, language proficiency, interaction with other colleagues

in the field, among others. In addition, the ADLAB PRO questionnaires were created after a comprehensive review of AD literature, AD guidelines and literature on questionnaire development, to name a few. Finally, the ADLAB PRO questionnaire (in fact, it is just one questionnaire that consists of three sections (excluding the demographic questions) that are answered depending on what group of AD stakeholders you belong to) was drafted by numerous experts accumulating decades of AD-related experience. Therefore, it was considered to be a trustworthy option. However, and as we indicated and specified in the corresponding publications, some modifications were necessary to make the questionnaire suitable for the Chinese case.

Another important issue to point out regarding the distribution of the ADLAB PRO-adapted questionnaire catering end users is that, despite the shortcomings derived from also administering it while people were queuing to watch a movie, it was the only possible option to avoid missing representative individuals. In fact, in the ADLAB PRO project it was already observed that only educated people are able to use online tools to answer a questionnaire and we were aware of the fact that many AD users have a poor education background. Regarding the online version of this questionnaire, and like we did in the case of the audio describers, the anonymous responses mode was selected in the online survey maker Web Survey Creator.⁷

With respect to our reception study, and unlike Fernández-Torné and Matamala (2015), we only used one synthetic voice. We were initially open to testing more than one voice, but the time constraints forced us to make a selection without users. We sent emails to various TTS companies including Chinese voices in their catalogue and asked for their cooperation in our academic research. Only one Spanish company replied, but the voice offered to us was not even considered to be comparable to the voice offered for free by the Chinese company that helped us with the AD mix and that already had experience with TTS AD. We also asked them to listen to the other voice we had obtained and they, as native Chinese speakers, concurred saying that theirs was much better.

One last remark must be made: although our research is not methodologically innovative due to it being partially based on approaches from previous projects and studies (see Chapter 2, Chapter 3 and Chapter 5 for more specific information), we believe it is of the utmost importance to make use of already existing research methodologies and tools in

⁷ It must also be clarified that, for example, Google Forms is not available in China.

different contexts to confirm their validity. Moreover, “now that experimental research has started to be conducted in AVT, the need for replication studies becomes more pressing” (Díaz-Cintas & Szarkowska, 2020, p. 8). We are aware of the fact that replication is not especially attractive or valued (Szarkowska & Wasylczyk, 2018, p. 14) but AVT research needs to be revisited for the sake of the discipline, as some voices in TS have already demanded for a while (Saldanha & O’Brien, 2014; Olalla-Soler, 2019). In effect, researchers such as Flis *et al.* (2020, p. 42) have warned that “[m]odern science is facing a replication crisis”. Replication can enhance the likelihood that a finding is true (Mooresinghe, Khoury & Janssens, 2007 in Olalla-Soler, 2019; Flis *et al.*, 2020, p. 42), since most published research findings are false (Ioannidis, 2005). Yet, Flis *et al.* (2020, p. 43) also advise caution:

Replication may not produce the same result as the original study for a number of reasons. First, the results of the original study could be a false positive, i.e. they could show a finding that is not really there, known as Type I error (Field 2009). Second, the replication study could be a false negative, i.e. it does not find a result when the result is genuine (Type II error). As noted by the Open Science Collaboration (2015: 943), “even research of exemplary quality may have irreproducible empirical findings because of random or systematic error.”

Olalla-Soler (2019) notes that Open Science Collaboration (2015) performed exact replications of 100 empirical studies and while 87% of the original studies had given statistically significant results, only 36% of the replications did. What is more, in a survey distributed to 1,138 Psychology researchers (Fiedler & Schwarz, 2016 in Olalla-Soler, 2019), 47% of the respondents confessed having used doubtful practices. In the TS field, Olalla-Soler (2019) conducted a study with 52 participants, 46.2% of whom had tried to carry out a replication of an empirical study to help consolidate previous studies or to verify earlier results, among other reasons. Yet, only 23.4% of these replications had obtained the same results as in the original studies. These are just some examples that prove the need to replicate previous studies so that science can move ahead with the required guarantees of success.

1.5. Structure of the dissertation

This dissertation is presented as a compendium of publications. The following four papers, which are presented chronologically to illustrate the different steps taken in our investigation, make up the main body of this research.

Tor-Carroggio, I., & Casas-Tost, H. (2020). Who is currently audio describing in China? A study of Chinese audio describer profiles. *MonTI. Monografías de Traducción e Interpretación*, 12, 78–107.

This first paper constitutes Chapter 2. This article contextualises AD within the Chinese context by examining the legislation on MA, its history and current state. It also presents a questionnaire-based study with Chinese audio describers from the most active AD centres in Mainland China, i.e., Shanghai, Guangzhou, and Beijing. It aims at profiling the individuals that are currently audio describing in China, since it is not a recognized job and there is little training available.

This article is linked to the first main aim of this thesis: to describe the status quo of AD in Mainland China, both as an access service and an object of study, with special focus on the city of Shanghai. More precisely, the specific objectives a) and b) are achieved in this paper.

Tor-Carroggio, I. (2020). The customer is always right: Study on Chinese persons with sight loss' opinion on their experience with audio description. *Disability & Society*. doi: 10.1080/09687599.2020.1713727

This second paper makes up Chapter 3. This article quickly revises how people with disabilities have traditionally been viewed in China and evaluates the habits, needs and satisfaction of Chinese AD users towards the AD they currently have access to through two complementary questionnaires mostly distributed in the city of Shanghai.

This article is also linked to the first main aim of this thesis: to describe the current state of AD in Mainland China, both as an access service and an object of study, with special focus on the city of Shanghai. More concretely, the specific objective c) is fulfilled in this paper. Chapter 3 also touches on the second main objective, since a comparative with the European case is included. More precisely, it accomplishes its specific objective b).

Tor-Carroggio, I., & Vercauteren, G. (2020). When East meets West: A comparison of audio description guidelines in China and Europe. *HIKMA. Translation Studies Journal*.

The third paper is included in Chapter 4. Its aim is to present some unpublished Chinese in-house AD guidelines according to which Chinese audio describers are currently being trained. Due to guidelines being one of the most researched topics in the AD field in the West, and especially in Europe, a brief comparison with the European guidelines and standards is also presented for the purpose of highlighting both differences and similarities.

This article completes the accomplishment of the first main aim of this thesis: to describe the status quo of AD in Mainland China. More concretely, the specific objective d) is achieved in this paper. Chapter 4 also touches on the second main objective, since a comparison with the European case is included. More precisely, it accomplishes its specific objective c).

Tor-Carroggio, I. (2020). T(ime) T(o) S(tart) synthesising AD in China? Results of a reception study. *JoSTrans: The Journal of Specialised Translation*.

The fourth and last article is included in Chapter 5. This paper discusses two main aspects: first, what type of voice Chinese AD users prefer (human or synthetic) and, second, whether TTS AD can be a solution worth considering and implementing to produce AD in a cheaper and quicker way both provisionally and in the long run. This article is based on the results of a reception study carried out in Shanghai with people with sight loss. The reception study consisted of both a quantitative and a qualitative part to yield a deeper insight into the possibility of increasing the offer of audio described films by opting for this alternative.

This article is exclusively linked to the third and last main aim of this thesis, namely, to determine whether TTS AD can be accepted in movies in China, and it achieves its two specific objectives.

All articles are reproduced as chapters and, as requested in all dissertations by compendium of publications at UAB, no changes were introduced, neither in terms of content nor format.

Additionally, Chapter 6 includes the part of the discussion that could not be added in the papers due to length limitations. It mainly draws comparisons between Chinese and European audio describers to fulfil the specific objective a) from the second main aim. At the same time, this chapter compiles the global conclusions reached after completing the whole process of research and achieving each of the objectives established at the beginning of the thesis. It also lists the limitations of this research and indicates some other avenues for future work that merit some attention. Finally, it also highlights the main dissemination activities carried out because we believe sharing our results and acquired knowledge was of utmost importance in China given the current state of AD in the country. This was also part of the AR approach that we have always been committed to adopting.

Moreover, there is a chapter (Chapter 7) listing all the bibliographical references consulted while drafting this thesis. It includes not only the references in the articles, but also those cited in the non-published parts of our thesis. Finally, the annexes including all complementary documents and materials used in the three studies have been added, namely:

- a) Articles included in this dissertation (Annex 1): the published versions of the articles presented in each chapter.
- b) Additional articles (Annex 2): complementary publications that, although not strictly related to this thesis, also revolve around China and AD. More specifically, the following articles have been included:
 - a. Liu, Y., & Tor-Carroggio, I. (2022). *Comer, beber, amar... y comparar. Análisis contrastivo de la audiodescripción en chino y español: un estudio de caso [Eat Drink Man Woman... A Contrastive Analysis of an AD in Chinese and in Spanish: A Case Study]. Onomázein.*
 - b. Xiao, W. (肖维青), & Tor-Carroggio, I. (2020). Shiting fanyi xin fazhan: yuyin hecheng zai koushu yingxiang zhong de yingyong. Jiyu zhendui shizhang renshi de jieshou shiyan yu diaocha (视听翻译新发展: 语音合成在口述影像中的应用——基于针对视障人士的接受实验与调查) [New developments in audio-visual translation: The use of synthetic voices in audio description. An acceptance experiment for people with

sight loss]. *Dongfang Fanyi* (东方翻译) [East Journal of Translation], 64(2), 24–30.

- c) The documentation submitted to UAB’s Ethics Committee and their approval for the study with audio describers, with users, and for the reception study. This annex includes the informed consent forms for the study with audio describers, with users and for the reception study as well (in both English and Chinese) (Annex 3).
- d) Questionnaires used in the study with Chinese audio describers and AD users, as well as those used for the reception study (in both English and Chinese) (Annex 4).
- e) AD script of the clips used in the reception study (Annex 5).
- f) Interviews with the author of this thesis that prove the impact our research has had in the Chinese society and the interest it has generated (Annex 6).

The inclusion of these interviews is especially relevant because they illustrate that, apart from the academic contribution this thesis has resulted in, efforts have already been made to transfer the acquired knowledge to society so that our work’s impact is not restricted to the academic world.

As requested by the Chinese associations and institutions that provided their AD guidelines, these materials are not attached.

Chapter 2. Article 1: Chinese audio describers' profile

Who is currently audio describing in China? A study of Chinese audio describer profiles⁸

Resumen

El acceso a la cultura y la información está reconocido por la Convención sobre los Derechos de las Personas con Discapacidad de las Naciones Unidas, ratificada por China en 2008. Los servicios de accesibilidad sensorial, tales como la audiodescripción, facilitan el consumo de productos audiovisuales a las personas con pérdida de visión. Este artículo contextualiza dicho servicio en China y se centra en aquellas personas a cargo de su preparación: los audiodescriptores. Para conocer su perfil, se distribuyó un cuestionario en Shanghai, Beijing y Guangzhou. Los resultados muestran que prácticamente todos los informantes son voluntarios de entre 20 y 50 años, con formación universitaria variada no relacionada con la Traducción Audiovisual. Ejercen profesiones muy diversas y apenas han recibido formación en audiodescripción, ámbito en el que tampoco tienen una trayectoria dilatada. Siempre audiodescriben en chino estándar y el cine constituye el producto audiodescrito por excelencia.

Abstract

Access to culture and information is recognized by the United Nations Convention on the Rights of Persons with Disabilities, ratified by China in 2008. Access services like audio description facilitate the consumption of audio-visual products such as films by those who suffer from sight loss. This paper attempts to contextualize this access service in China and focuses on those in charge of preparing it – audio describers. In order to provide a profile of these audio describers, a questionnaire was distributed in Shanghai, Beijing and Guangzhou. The results show that an overwhelming majority of those audio describing are volunteers aged 20-50, and most have a university education background which is not related to Audiovisual Translation. They come from a varied professional background,

⁸ Tor-Carroggio, I., & Casas-Tost, H. (2020). Who is currently audio describing in China? A study of Chinese audio describer profiles. *MonTI. Monografías de Traducción e Interpretación*, 12, 78–107.

most have only been audio describing for a few years, they lack formal training in audio description and mainly audio describe films in Standard Chinese.

Palabras clave: Accesibilidad a los Medios. Servicios de Accesibilidad. Audiodescripción. Audiodescriptores chinos. Voluntarios.

Keywords: Media Accessibility. Access Services. Audio Description. Chinese Audio Describers. Volunteers.

2.1. Introduction

According to the United Nations, China is the country which has the highest number of citizens with disabilities in the world (cf. Qiu 1998; Wu & Xie 2015: 21). The most recent survey, the Second National Sample Survey on Disability (2006), revealed that there are almost 83 million persons with disabilities in China, 6.34% of the country's population. In more specific functionally diverse terms, 29.07% have a physical disability, 24.16% suffer from hearing loss, 14.86% have sight loss, 7.40% are mentally disabled, 6.68% are intellectually challenged, 1.53% have a speech disability, while 16.30% suffer from multiple disabilities. The survey also pointed out that more than half of the respondents (53.24%) are at least 60 years old and that the vast majority of them (75.04%) live in rural areas, where residents frequently lack adequate access to healthcare, let alone to culture or information. The data from 2006 showed an increase in the number of persons with disabilities compared to the results of the first official survey in 1987, which can be attributed to, among other reasons, China being one of the societies with the highest population age, which, in turn, is closely associated with disability (cf. Peng et al. 2010).

Although China signed the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) in 2007 and ratified it one year later, it still has a long way to go before it can guarantee and implement the rights of the functionally diverse. Access to culture is one of the rights the UNCRPD recognizes, which includes "access to television programmes, films, theatre and other cultural activities." This can be accomplished, though not exclusively, through sensorial accessibility services such as audio description (AD), subtitling for persons with hearing loss and sign language interpreting. Of these, AD is the focus of this article and has been defined by Jankowska (2015: 9) as "an

additional verbal description of the most important elements of a picture, i.e. a technique that provides persons with sight loss with information that other viewers perceive only visually". It has also been described as "a verbal commentary providing visual information for those unable to perceive it themselves" (cf. Fryer 2016: 1), or simply as "the visual made verbal" (cf. Snyder 2014). Although mainly targeted at persons with sight loss, many studies underline its potential for immigrants or language students, so in fact the actual scope of AD is quite wide-ranging and not just restricted to those with disabilities (cf. Gambier 2003; Snyder 2014; Walczak 2017; Di Giovanni 2018).

The academic study of AD has been embraced by Translation Studies, more specifically Audiovisual Translation (AVT) (Gambier 2003), alongside the study of, for instance, subtitles for persons with hearing loss. In fact, a new subfield has emerged within AVT, called "media accessibility", which specializes in the study of sensorial accessibility services. The reason for including AD under the Translation Studies umbrella lies in "AD implying a semiotic rendering of images into words" (cf. Hernández-Bartolomé & Mendiluce-Cabrera 2009: 5). This intersemiotic translation involves any kind of nonverbal information, for example, costumes, gestures, colors, body language and sounds, which have to be transmitted by means of words.

Officially born in the US in the 1980s (cf. Fryer 2016: 15), AD is an accessibility service currently available in different parts of the world, albeit to differing degrees, which include not only many Western countries, but also those in Asia, such as Japan and Thailand.⁹ In countries like Spain, Poland and the United Kingdom, there is even specific legislation that requires broadcasters to offer a particular amount of accessible content on TV, part of which has to be audio described. In the same vein, some countries such as Spain, France, Belgium, Greece, Italy and Poland have drafted their own AD standards or guidelines to "describe the best practices that should be applied for quality AD services" (Reviere 2016: 236). China has numbered among countries where AD is a reality since the early 2000s, however, as Wu and Xie (2015: 21) revealed, the general public are largely unaware of it. AD is still in its infancy in China for various reasons: its dependence on volunteers at grassroots level, the fact that movies are the only field where it is used and the lack of official standards and supporting legislation in the country. Another fact

⁹ See the Media Accessibility Platform's Accessometer for a world map of the legislation, standards and guidelines on media accessibility organised by countries: <https://mapaccess.uab.cat/accessometer>, accessed on 9 January 2019.

which illustrates its early stage of development as a service in China is the array of terms used to refer to it: “barrier-free movies” (*wuzhang'ai dianying* 无障碍电影) -in which theoretically subtitles for persons with hearing loss and sign language interpretation are also added-; “commentary” (*jieshuo* 解说) and “image description” (*koushu yingxiang* 口述影像). The latter has already been taken up by some researchers who have already started using this term in academic publications, including Leung (personal communication), who advocates its use for the purposes of consistency in the academic world.

Similarly, unlike Europe, where the study of AD has already come of age (Reviere 2016), academic AD-related research in China is still in its infancy as illustrated by the dearth of references to it in bibliographies (cf. Gambier & Jin 2018). Research into the development of AD in Hong Kong has been carried out by Yeung (2007), and later Leung (2015) who describes its history and lists the domains where it is used nowadays (museum exhibitions, films, outdoor activities, etc.), noting that it is more widespread than in Mainland China. Nevertheless, both researchers admit that Hong Kong is lagging behind some Western countries because it still lacks a regulatory framework for providing AD that would make it comparable. Regarding Taiwan, the leading figure in research into AD is Chao Ya-li (赵雅丽), the first researcher to carry out empirical studies on AD in Greater China. However, in recent years academic research into AD in Taiwan has received very little attention. In the case of Mainland China, one of the first studies into AD was Li (2013) and later Leung (2015), both of whom pointed out that it was restricted to films only, though not always available. Li and Pan (2013) reported on the early years of AD in Mainland China, especially in Shanghai, and enumerated some of the challenges the service still has to face, such as the lack of financial resources and audio-visual material, the poor technical level and its uneven provision across the country. Although Feng (2018) set out the various fields in which AD can be used, AD still continues to be relegated to films only on the Mainland.

While the trend in the West is to carry out reception studies that attempt to delve into user preferences, needs and characteristics (cf. Chmiel & Mazur 2012; Fresno 2014; Matamala et al. 2018), Chinese AD is a few steps behind since it still lacks a well-defined descriptive framework that contextualises the service and provides a clear picture of how it is offered. Given the scarcity of literature on AD in China, especially in Mainland China, the

objective of this study is twofold. First, to investigate the history and the state of the art of this accessibility service in China, including the corresponding legislation. Second, to provide an accurate profile of audio describers in Mainland China, i.e. the characteristics of the key agents currently working on ensuring that Chinese with sight loss have access to culture as a human right.

This article is divided into four sections. The first describes the methodology, while the second contextualizes AD within the Chinese context by examining the legislation on media accessibility, its history and current state. The third analyzes the profile of Chinese audio describers based on the results obtained from a survey distributed among the main agents in the most active AD centers in Mainland China, Shanghai, Guangzhou, and Beijing. The final section offers conclusions and proposals for future improvements.

2.2. Methodology

The first part of this research is based on a thorough and extensive review of the literature published on this subject and is complemented with interviews with key AD agents who helped us to map out a more precise picture of the current state of AD as well as of the profile of audio describers in particular. These include the following interviewees:

- a) Qu Dapeng (曲大鹏): Head of the volunteer team of audio describers from 990 News, interviewed at Cathay Cinema (Shanghai) on June 29, 2017.
- b) Han Ying (韩颖): Founder of Sound of Light Barrier-free Film & TV Culture Development Center, interviewed in her office in Shanghai on July 10, 2017.
- c) Jiang Hongyuan (蒋鸿源): first audio describer in Shanghai, interviewed at Cathay Cinema (Shanghai) on September 28, 2017.
- d) Huang Yiqing (黄一庆): General Secretary of the Shanghai Film Critics Society, interviewed in his office on July 3, 2017.
- e) Dawning Leung (梁凯程): Founder and CEO of the Audio Description Association Hong Kong, interviewed in her office in Hong Kong on August 10, 2018.
- f) Luo Yisheng (罗逸生): AD sessions manager at the Guangzhou Library, interviewed through Wechat on December 11, 2018.

- g) Yao Weiming (姚伟明): Head of Yangguang Volunteer Association in Wuxi (Jiangsu Province), interviewed through Wechat on December 11, 2018.
- h) Qin Xiaojie (秦晓婕): Head of AD activities in the Sun Yat-sen Library of Guangdong Province, interviewed through Wechat on January 29, 2018.

The profile of Chinese audio describers was drawn from the information gathered through a survey, which was taken from the EU-funded project ADLAB PRO.¹⁰ Although the project was circumscribed to Western countries, the questions were also appropriate to illustrate Chinese audio describers' profiles and their experience with AD. As one could have predicted, the questionnaire did not completely fit the Chinese profile and, with the survey creators' consent, had to be adapted. The following are the seven major changes that were applied. First, the original version asked respondents which country they were from, but, since our informants would all be Chinese, we asked from which province instead. Second, rather than maintaining the original question of what their mother tongue was, only two options were finally given: "Standard Chinese" and "Chinese dialects" (for which respondents could specify). Third, if the respondents answered "no" to the question asking if they had ever collaborated with anyone who has sight loss when preparing audio descriptions, it was deemed necessary to add whether they thought cooperating with them would have a positive influence on their work. Fourth, the option "I am a student" was added to the question about the job audio describers had, since in China many of them are still university students. Fifth, due to the lack of official standards or even national guidelines on AD, a question was added to find out the opinion of those currently audio describing in China on the need to have such standards in their country.¹¹ Sixth, the option "I do not have a specialization" to the question aimed at finding out whether respondents always accepted all AD jobs, even if they fell outside their area of specialization. The reason for this is because, to the best of our knowledge, you only find AD applied to movies, so it would be difficult to come across any other field of AD specialization. Finally, the original version of the questionnaire often uses the collocation "write audio descriptions", which was considered inaccurate here as Chinese audio describers usually

¹⁰ See <https://www.adlabpro.eu/>, last accessed in May 2017.

¹¹ There is an international standard on AD issued by the International Organisation for Standardisation (TS 20071-21 "Information Technology — User interface component accessibility — Part 21: Guidance on audio description"), but to the best of our knowledge, it is still not a reference in China.

either write or voice AD scripts, but the profession given is the same in both cases. Therefore, the verb “prepare” in Chinese was preferred in most of the questions.

2.3. Audio description in China

This section describes AD in China from two perspectives. First, a brief overview of legislation setting out current media accessibility both within the Chinese and international legal frameworks. Second, a summary of the early stages of AD in Mainland China since the first audio described movie, as well as the art of AD in four major Chinese cities.

2.3.1. Chinese legislation on media accessibility

A country's legislation reflects its priorities and the importance it attaches to different issues, and China is no exception. Their legislation addressing the rights of citizens with disabilities has been developed in both quantitative and qualitative terms, illustrating China's concern to guarantee the equal rights of all its citizens.

Article 45 of the Chinese Constitution (1982, last amended in 2018) mentions persons with disabilities whose rights to material assistance “from the state and society” are recognized on par with the elderly and the ill. The Constitution is also committed to helping “make arrangements for the work, livelihood and education of the blind, deaf-mutes and other handicapped citizens.” Yet, the legal document that has resulted in the most significant changes in China is the afore-mentioned UNCRPD, which was approved in 2006. Its aim is “to promote, protect and ensure the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities, and to promote respect for their inherent dignity.” Article 30 of the UNCRPD serves as a legal framework for AD, as well as for other sensorial accessibility services, since it recognizes the right of persons with disabilities to take part on an equal basis with others in cultural life. It also requires member states to take appropriate measures to ensure that persons with disabilities enjoy access to:

- a) cultural materials in accessible formats;
- b) television programs, films, theatre and other cultural activities, in accessible formats;

- c) places for cultural performances or services, such as theatres, museums, cinemas, libraries and tourism services, and, as far as possible, enjoy access to monuments and sites of national cultural importance.

China, and subsequently Hong Kong and Macao, signed it in 2007 and ratified it one year later.¹² This ratification brought about several changes in the Chinese legislation, such as the revision of the Law of Protection of Persons with Disabilities (*Zhonghua Renmin Gongheguo Canjiren Baozhangfa* 中华人民共和国残疾人保障法), which had been passed in 1990, because it conflicted with the UNCRPD. Petersen (2008: 620) reminds us that the first version of this law was much more patronizing and clearly framed within the medical model of disability.¹³ In the first article of the current version, the objective of the Law of Protection of Persons with Disabilities is to protect the rights and interests of persons with disabilities by ensuring their “equal and full participation in social life and their share of the material and cultural wealth of society.” It covers several areas, one of which is cultural life, which has to be enriched by “offering TV programs in sign language and radio programs specially designed for disabled persons and providing subtitles or narrations for more TV programs and movies.” This is the first mention of AD in a legal document in Chinese history, albeit a somewhat vague reference. To date, both Taiwan and Hong Kong lack legislation that explicitly mentions AD or any aural narrations at all, although it must be stated that in actual practice AD can be found in more fields than on Mainland China, especially in Hong Kong.

There is national -and rather limited- legislation that defends the provision of sensorial accessibility services, however each province applies them (or not) depending on its socioeconomic possibilities. For example, Shanghai drafted its Measures for the Implementation of the Law of the People's Republic of China on the Protection of Disabled Persons (*Shanghai shi Shishi Zhonghua Renmin Gongheguo Canjiren Baozhangfa Banfa* 上海市实施中华人民共和国残疾人保障法办法) in 2013, which stipulates that the cinemas with the necessary means shall offer the so-called “barrier-free movies”. Likewise, the Shanghainese government is committed to offering a wide range

¹² Taiwan has not been given the chance to sign the UNCRPD but, as Rovira-Esteva and Zheng (2017) pointed out, it has drafted its own legislation allowing this territory to be at the level of other developed countries in accessibility matters, such as the Act to Implement the Convention on the Rights of Persons with Disabilities (Shenxin Zhang'aizhe Quanli Gongyue Shixingfa 身心障碍者权利公约施行法) in 2014.

¹³ According to this model, disability is a pathological and permanent issue that exclusively belongs to the individual, and that needs to be addressed.

of accessible cultural services for free or with special discounts, embodied in the 13th Five-Year Plan Development Plan for the City of Shanghai (*Shanghai shi Canjiren Shiye Shisan Wu Fazhan Guihua* 上海市残疾人事业十三五发展规划). It is also committed to setting up a venue in each district where those with sight loss can watch audio described movies. Beijing also drafted its own measures to guarantee the enforcement of the aforementioned law in 2011, but it does not specifically mention “barrier-free movies”. Instead, it sticks to the “narrations” on TV the national law briefly touches on, without giving any further details.

In sum, the Chinese national legislation already incorporates a term that can be related to AD -“narrations”-, although it is only conceived of in the domain of TV and its definition proves elusive or vague: many of the AD providers interviewed did not seem to know what this term means exactly. This lack of definition can also be seen in some implementation measures that have been drafted in each administrative region, such as in the case of Beijing. Shanghai has shown more commitment but, surprisingly it does not offer it on TV and along with the cities at the forefront of providing AD in China, the service is limited to films shown in cinemas. Finally, other provinces such as Guangdong have omitted the term “narrations” and only encourage the progressive inclusion of subtitles and sign language interpreting in TV content.¹⁴ So it can be seen that the legislative context for AD is quite diverse across the country, underlining how far China still has to go.

2.3.2. History and the state of the art of AD in Mainland China

Following on from this review of AD legislation in China, this section presents a brief history of the actual provision of this service around the country and the current state of AD in some cities which have succeeded in offering it on a regular basis. These are not the only cases, but they do constitute a representative sample of different associations or entities that share a common goal: to guarantee one of the rights of Chinese citizens with

¹⁴ “广东省实施《中华人民共和国残疾人保障法》办法(2018 修订)” (Guangdong shen Shishi “Zhonghua Renmin Gongheguo Canjiren Baozhangfa Banfa” 2018 xiuding, Guangdong’s Measures to Implement the Law on the Protection of Persons’ with Disabilities, amended in 2018), <http://www.csicare.com/Law/Show?id=223823> (accessed on 4 February 2019).

sight loss as granted them by the Chinese government when it ratified the UNCRPD, namely, their right to culture and information.

AD officially appeared in Mainland China on 23 April 2009 with Jiang Haiyang's movie *Examination 1977* (*Gaokao 1977* 高考 1977), which was screened with live AD at the Shanghai Library (*Shanghai Tushuguan* 上海图书馆) to an audience of 400. The film's AD was even voiced by a famous actor and the event gave birth to the Free-Barrier Movie Workshop (*Wuzhang 'ai Dianying Gongzuoshi* 无障碍电影工作室), made possible thanks to the cooperation between the Shanghai Association of Persons with Disabilities (*Shanghai shi Canlian* 上海市残联), the Shanghai Film Critics Society (*Shanghai Dianying Pinlin Xuehui* 上海电影评论学会) and the Shanghai Library. The workshop focused on the production of CDs of audio described movies that were later distributed among the associations for persons with disabilities, the libraries of each district and even donated to other provinces. The workshop did not have its own facilities, and it still does not; instead it relies on other companies allowing them to use their equipment for free.

Examination 1977 was not the first movie to be audio described, but it can be regarded as the culminating point in the first stage of AD in China, since one of its pioneers in China, Jiang Hongyuan, had devised this accessibility service in 2006. Jiang Hongyuan, who suffers from sight loss himself and used to work in the Chinese film industry from which he is now retired, claims to have come up with the idea from scratch after realizing how interested those with sight loss were in watching movies. His idea was conceived as early as 2002 but the initial reticence by associations hindered the provision of the service. He later managed to gather volunteers – some of whom also suffered from sight loss – and started recording AD for movies in CD format. From 2011 onwards, the service spread around China alongside the free-barrier movies being included in the 12th five-year plan (2011-2015) of the General Administration of Press and Publication (*Zhongguo Renmin Gongheguo Xinwen Chuban Zongshu* 中华人民共和国新闻出版总署) as a key project. As of 2012, on the last Thursday of every month a movie is audio described in a live session in one cinema of each district of the city of Shanghai. This is possible thanks to Jiang Hongyuan's monthly AD script and also to the collaboration of professional radio presenters from the 990 News (*990 Xinwen* 990 新闻) who volunteer to voice the scripts in the 17 cinemas that offer this service. This monthly activity is funded by the Shanghai

Association of Persons with Disabilities, which covers the expenses for the tickets of all the movie-goers. As mentioned earlier, the Chinese audio describer comprises two different figures: the scriptwriter and the voicer. The voicers are all enrolled in a voluntary project that started with only 8 volunteers and which now includes more than 300 radio presenters who share the work every month among themselves according to their availability through a Wechat group. Jiang Hongyuan also records movies with AD on CDs, which are then distributed among cultural centers for the elderly. Although the challenges to be tackled are still plentiful, there is one that stands out above the rest: how to deal with film copyright protection. Some AD providers who record CDs without purchasing the film's copyright turn to their own interpretation of the law to justify what they do. They argue that there is no need to buy the rights of any movie that is to be audio described because they extrapolate to films what article 22 of the Copyright Law of the People's Republic of China states:

In the following cases, a work may be used without permission from, and without payment or remuneration to, the copyright owner, provided that the name of the author and the title of the work are mentioned and the other rights enjoyed by the copyright owner in accordance with this Law are not prejudiced:

[...]

(12) transliteration of a published work into braille for publication (People's Republic of China, 2001).

Still in Shanghai, Sound of Light (*Guanying zhi Sheng* 光影之声) is an association led by Han Ying and set up in 2016, and it receives material and financial assistance from the Shanghai Association of Persons with Disabilities. They produce about 50 audio described movies every year, which are recorded in their facilities and distributed, unlike Jiang Hongyuan's team, through digital means for ecological and copyright reasons. These movies can be accessed at specific facilities, which are all located in Shanghai. The movies are likewise audio described by volunteers, including the same radio presenters that voice the monthly AD live sessions in cinemas. Scriptwriters, who tend to be Sinology or Communication university students, need to sign a copyright protection agreement that prohibits them from posting their scripts on the Internet. After passing a test before being accepted in the association, they are also offered a couple of training sessions and when they start their first project they are always guided and supervised personally by Han Ying, who is also an AD consumer herself.

Although Shanghai is the city in which AD has developed the fastest, Li (2013: 145) states that there were already movies with AD in Beijing as early as 2005, which were shown in rather poorly equipped screening rooms thanks, again, to the help of volunteers (cf. Li 2013: 145). Beijing is also the city where the China Braille Library is located, which started screening films with AD in 2011, the first of which was Han Sanping and Huang Jianxin's *The Founding of a Republic* (建国大业) (cf. Xu 2018). This library is also the setting for the Kangyi Barrier-Free Movies Development Center (*Kanyi Wuzhang' ai Dianying Yingshi Fazhan Zhongxin* 康艺无障碍影视发展中心). This center is the only one selling CDs of audio described movies to the whole country, which retail for around RMB10 each. Yet, their annual production is rather low and the movies are also quite dated because only CDs with purchased copyright are marketed.

Another of the major Chinese cities offering AD is Guangzhou. The Sun Yat-sen Library of Guangdong Province (*Guangdong Shengli Zhongshan Tushuguan* 广东省立中山图书馆) has been offering live AD sessions every month since 2007. It also lets users borrow movies that already incorporate AD and that are bought from the China Braille Library. Nonetheless, and according to some of our interviewees, users prefer to attend live sessions because the movies screened are newer and more appealing to them. Most of the audio describers are, again, volunteers recruited through announcements in the Library. The Sun Yat-sen Library of Guangdong Province even offers AD training and assists other AD providers inside and outside their province, such as the Guangdong Library (*Guangdong Tushuguan* 广东图书馆).

Apart from first-tier cities, other places like Wuxi also offer AD. The Wuxi Yangguang Volunteer Association (*Wuxi Xinwuqu Yangguang Zhiyuanzhe Xiehui* 无锡新吴区阳光志愿者协会) uploads audio described movies to Youku and other audio sharing platforms such as Ximalaya FM (喜马拉雅 FM), which provides audiobooks, music, a variety entertainment programs, news, and other audio contents. In many cases, they upload the AD track and the movie sound track, but not the images. Ximalaya FM is also home to other audio described movies posted by the Shanghai Film Critics Society and is one of the few channels AD users can access from home.

Apart from these examples, and according to Wu and Xie (2015), in recent years the government has supported AD and movie-related initiatives in the provinces of

Heilongjiang, Chongqing, Liaoning, Jilin, Hunan, Henan and Ningxia, to name but a few. In other words, there is a short history for providing AD in China, it has been focused on a handful of big cities, but it has recently expanded to other areas.

Before administering the questionnaire, and given the increasing necessity to apply for ethical approval in AVT research (cf. Orero et al. 2018), the questionnaire was first submitted to our university's Ethics Committee for approval. It was first piloted with two audio describers and then distributed with the help of cooperating associations. The survey was administered through Web Survey Creator and was open from July 2017 to December 2018. We obtained valid responses from 53 informants, who stated they were part of Shanghai, Beijing and Guangzhou audio describers' teams. The results were analyzed statistically using IBM SPSS (version 22) and are presented in the results section.

2.4. Profiling audio describers: Results of the questionnaires

In order to properly profile Chinese audio describers, they were sent a survey to obtain first-hand data that would complement the historical and legislative dimensions presented so far.

This questionnaire had four sections: first, demographic information; second, details about the informants' activity as audio describers; third, information about the status of AD and their professional circle and, finally, information about their educational background and training specific to AD. The original version of the questionnaire also included a fifth section focusing on the competences and skills of audio describers, the results of which were finally discarded for this paper, since the objective was to have a general picture of those audio describing, rather than details of their day-to-day activity.

Section one of the questionnaire focused on the demographic profile of the respondents. Most of the 53 participants of the survey prepared AD in Shanghai, although some of them also audio described in Beijing (12 respondents), and Guangzhou (five respondents). Yet, their place of origin is more varied: although half of them are from Shanghai (52.8%), many come from other provinces, such as Guangdong (11.3%), Jilin (5.7%), Anhui (5.7%), Hebei (3.8%), Heilongjiang (3.8%), the municipality of Beijing (3.8%), Inner Mongolia (1.9%), Shandong (1.9%), Xinjiang (1.9%), Henan (1.9%), Liaoning (1.9%), Hubei (1.8%) and Zhejiang (1.9%). Most of the audio describers were women (75.5%),

almost all the respondents were between 20 and 50 years old (96.2%) and almost half between 31 and 40 years old. More than 80% of the audio describers stated Standard Chinese was their mother tongue, 20% specified other dialects, but almost all of them prepared AD in Standard Chinese (98.1%). As for their educational background, all the informants had at least a bachelor's degree, except two: one had attended a vocational school while the other had only completed secondary education. Allowing for one exception, the informants stated they did not suffer from sight loss (98.1%).

Section two of the questionnaire included a battery of questions regarding the informants' activity as audio describers. With the exception of two, who stated they had been audio describing for more than ten years, 80% of the audio describers had been offering this accessibility service for five years at most, with 34% of them having done so for less than a year. The survey also confirmed that the vast majority audio describing (almost 71.7%) were volunteers who would not charge any fee for their services. Yet, it cannot be stated that all audio describers in China do it on a voluntary basis since 17% of the respondents admitted being paid for their work. The same percentage stated AD was a semi-professional job for them and, instead of money, they received a token payment or payment in kind. For this question choosing more than one answer was permitted.

64.2% of the respondents admitted to not having had any AD-related training, while 35.8% had been taught how to audio describe. Half of those who had been trained had attended training lessons conducted by a company or an association. Unfortunately, out of all the audio describers who received training, as many as 77.6% had not obtained a certificate and 95.9% of all those surveyed acknowledged that they had never been required to present any form of certificate to work as audio describers.

When asked about the fields in which they had received training (a question which permitted multiple answers), it is cause for concern that almost 40% answered not having received any training at all, as shown in Figure 2.1. Despite the fact that almost all the respondents (94.3%) admitted the product they audio described was mostly films, those cases where they had received some training were courses which did not specialize in this particular field: 24.5%, as Figure 2.1. shows.

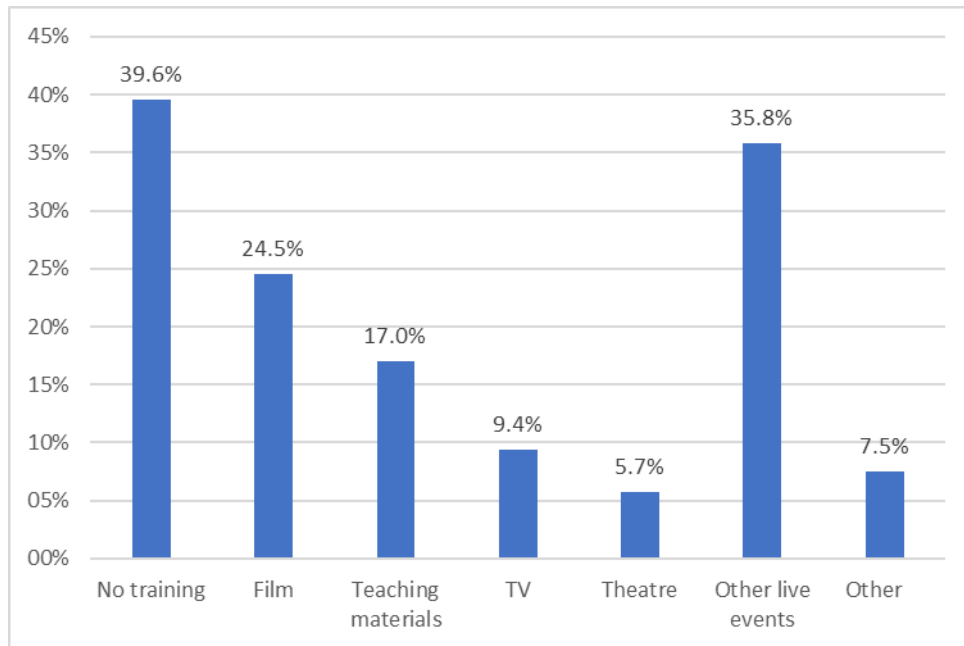


Figure 2.1. Fields in which audio describers have been trained

Another relevant conclusion is that audio describers in China do not have wide experience in audio describing. Taking into account how much material the respondents had produced in their career as audio describers, and that many of them were volunteers that had been preparing AD for a relatively short period of time, it comes as no surprise that 62.3% of them had audio described for less than 50 hours (Figure 2.2.).

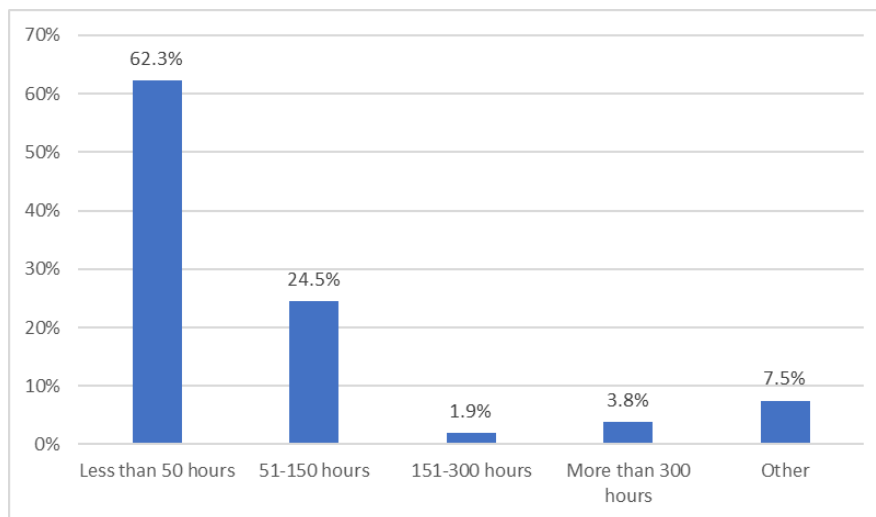


Figure 2.2. Audio describers' experience in hours

When asked about the specific tasks they performed, one can conclude that Chinese audio describers are assigned different tasks: voicing the AD scripts (67.9%) written by the scriptwriters (41.5%) the most popular one in our sample. More than a quarter of those

surveyed (34%) applied their voice talents to help record AD; 17% mixed the AD with the original soundtrack; 11.3% supervised the quality of the product and 3.8% were in charge of other AD-related tasks.

As for their working habits, Figure 2.3. shows that there is no clear pattern about whether they work alone or within a team, although the two options of always doing it in a specific way ranked lowest.

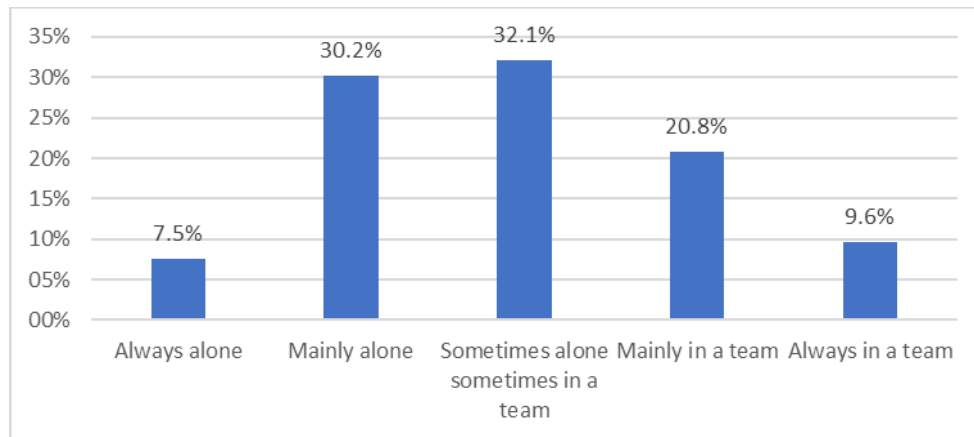


Figure 2.3. Frequency of collaborative work

Speaking of collaboration, almost half of the respondents either always or often cooperated with persons with sight loss during the creation of AD (Figure 2.4.).

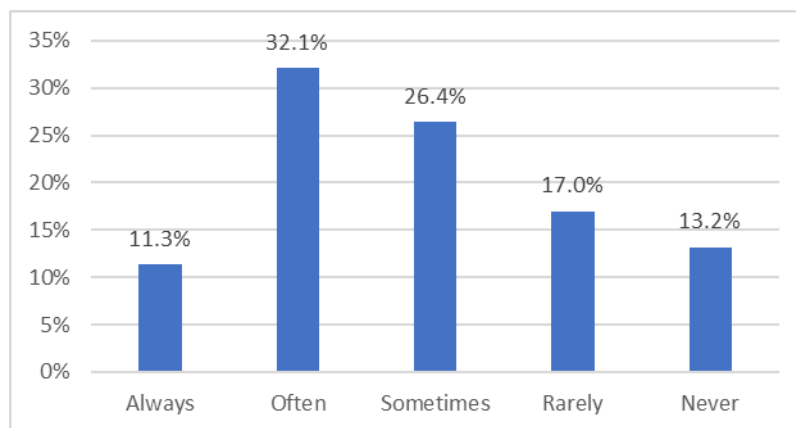


Figure 2.4. Frequency of collaboration with persons with sight loss

When asked about whether they thought they could benefit from working with persons with sight loss if they had never done so, 69.8% said yes, while 11.3% did not agree. 18.9% did not answer the question, probably because they did not have – or had little – experience. Thus, it can be seen that more than half of the respondents who have never worked with a person with sight loss when preparing an AD thought it would be helpful

to do so. This is a very positive fact because it shows that audio describers are aware of the importance of working with end users, even if they have never done so before.

Section three of the questionnaire evaluated the status of AD and the professional circle. Almost all the informants (94.3%) had previous working experience, many of them being presenters (42.0%), teachers (18.0%), journalists (14.0%), radio/TV commentators (4.0%), actors (4.0%), TV producers (4.0%), writers (2.0%) and others (38.0%). It is worth highlighting that nobody had ever worked as a translator before, let alone as an audio-visual translator, which is the field where Western audio describers are trained in. This finding is not surprising taking into account that media accessibility is not a university subject in China yet, not even in Translation Studies. Only 15.1% of the respondents worked in the AD field full-time, which is in line with the percentage of respondents working professionally, i.e. making a living, in the field (17%).

Those participating in this study were also asked about their gateway into AD and more than half of them got to know about it thanks to an acquaintance who was already working in the field (Figure 2.5.). This result makes sense taking into account that no specialized courses can be found at university level and those that exist are exclusively available to those who are already committed to offering this accessibility service. Even though AD has already appeared on the news and there are events organized from time to time to recruit volunteers, it is understandable that the most common way to get to know about this service is through somebody who is already involved in it. Also, one's CV is not what matters the most, and in many cases it is not even necessary to present it because it is generally assumed that the candidate has no previous experience in the field. It also needs to be highlighted that if audio describers are always recruited through word of mouth, this could result in the perpetuation of AD being in the hands of volunteers, and not becoming a professional job profile.

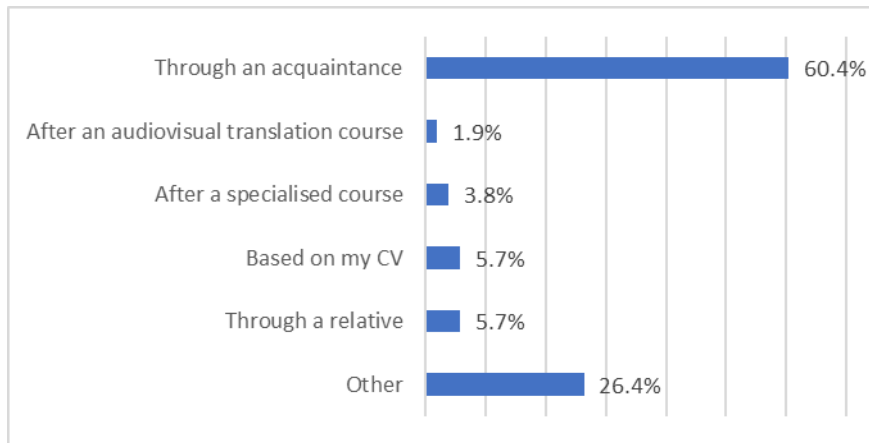


Figure 2.5. Gateway to AD

Three quarters of audio describers claimed to often (30.2%) or sometimes (43.4%) ask other audio describers their opinion on how to overcome certain AD issues they need to tackle. No respondent stated that they had never asked other audio describers' questions. As for reasonable deadlines to submit their work, 35.8% of the informants had always had enough time to finish the AD; 39.6% often had enough time; 18.9% sometimes had a deadline they felt comfortable with to satisfactorily audio describe, while only 5.7% of the participants rarely had sufficient time to finish the task adequately. This is not surprising as most work is on a voluntary basis.

Only 30.2% of those surveyed knew about existing AD guidelines, whereas 69.8% were completely unaware of it. From another angle, 83.0% thought guidelines were important; 9.4% did not think so and 7.5% thought that they were important to a certain extent. Likewise, we wanted to know if there was any link between knowing about guidelines and the importance attached to them. We realized that all respondents who stated they knew about guidelines (16 out of 53) considered them to be important. The association was evaluated and was found to be statistically significant at the alpha level of 0.1 but not at the 0.05 level, $\chi^2(2, N = 53) = 4.68, p < .05$.

When it comes to the audio describers' opinion on AD, and more precisely on whether it is an art or a craft, Figure 2.6. demonstrates that opinions are diverse, which does not allow us to draw any clear conclusions.

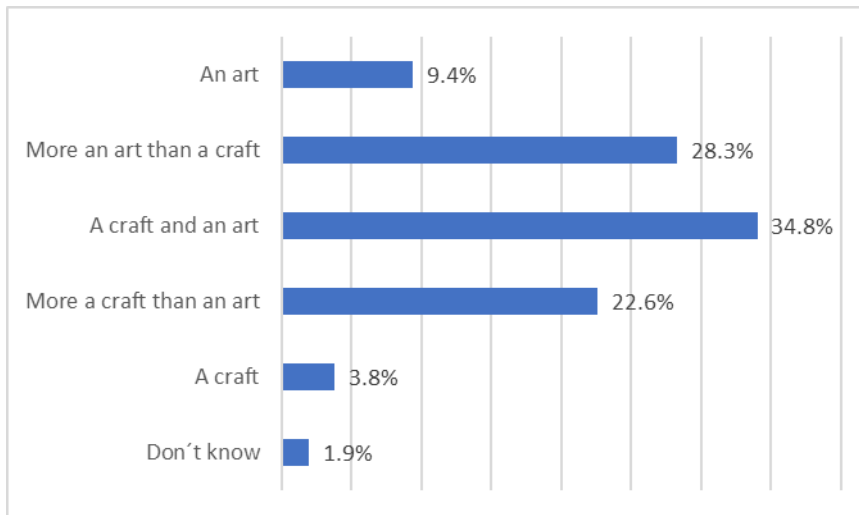


Figure 2.6. Notion of AD

The respondents' view on some statements related to AD is presented in Figure 2.7., for which they had to score on a one to five scale. On the one hand, it can be seen that the statement on which informants agree the most is the one that underlines the social impact of what they do. On the other hand, the statement that they disagree on the most with is the one related to AD being a well-paid. This makes sense taking into account that, as we have seen, in most cases financial compensation is non-existent. Also, and as expected, AD is not a well-known service in general, and this is probably why it is not considered a prestigious job.

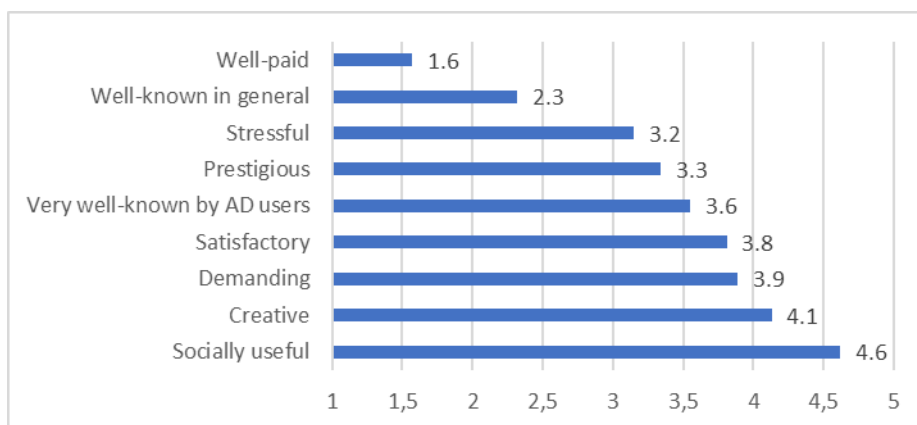


Figure 2.7. Perception of the qualities of being an audio describer

Section three ends by asking what figure an audio describer most resembled. As expected, and since many of our informants were radio commentators themselves, the most popular option was “presenter or commentator” (54.7%). It is not surprising that, unlike what

happens in the West, where they see audio describers as audiovisual translators,¹⁵ only 15.1% saw some kind of association between what audio-visual translators and audio describers do. This is probably due to the fact that China still does not include media accessibility in the Translation Studies curricula. According to our informants, other figures that had some points in common with audio describers were scriptwriters (18.9%), actors (7.5%) and writers (3.8%). Respondents could choose as many answers as they liked.

Section four focused on respondents' educational background and AD training. It starts by asking what field of knowledge they come from, which proved to be rather diverse. Figure 2.8. shows that more than 50% of the respondents had a Language/Linguistics or Journalism/Media Studies background. Only one informant had studied Translation.

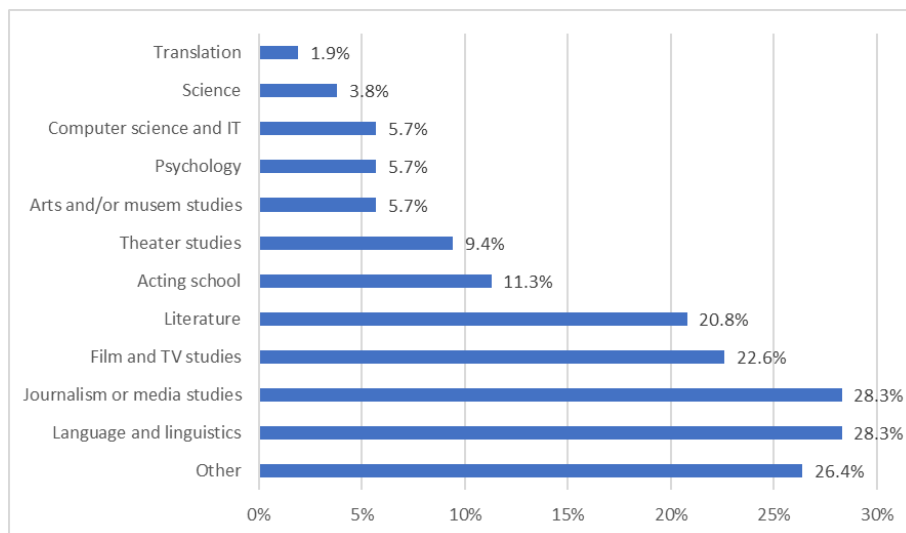


Figure 2.8. Audio describers' educational background

Since training is not the main gateway to AD, acquiring experience in the field is regarded as the main means by which audio describers have to continue to improve their skills and competences (84.9%); followed by attending conferences and workshops or analyzing existing ADs focusing on the solutions adopted by colleagues (both with 54.7%); studying existing material such as guidelines, academic articles, books on AD, (39.6%); doing research by gathering information on the product or by discussing with film

¹⁵ See the report on Audio description professional: Profile definition of the EU-funded project ADLAB PRO <https://www.adlabpro.eu/wp-content/uploads/2018/04/IO2-REPORT-Final.pdf>, last accessed in January 2018.

directors or producers is also an option (34%); and receiving in-house training (30.2%), while only 9.4% admitted not considering how to improve their skills at the moment.

2.5. Conclusions and proposals for the future

This article has shown that AD in Mainland China is still in its infancy, especially when compared to some countries in the West. Yet, cities such as Beijing, Shanghai and Guangzhou have launched some AD-related initiatives, all of which are related to films. This clearly manifests their interest in guaranteeing persons with sight loss access to information and culture and some of these initiatives have even been shared with other less developed regions in the country.

The situation is similar from an academic perspective: AD as an object of study is almost non-existent to date in China, with the exception of a handful of Chinese scholars who have elaborated some descriptive studies. Hong Kong and Taiwan must be considered separately because empirical research has already begun in these two regions, although the researchers doing so are extremely scarce. Of the many challenges that AD must face in China, and as Feng (2018) had already pointed out, copyright is the one that cries out for attention the most and the one which all AD providers coincide in mentioning and complaining about. The reason for this is that the current situation does not allow AD providers to use new films, unless they purchase the film copyright. The alternative involves audio describing them live, which ends up being the most widespread practice.

Audio describers are crucial to providing this accessibility service, but China lacks unified training and standards that ensure the quality of their work. Nevertheless, cities such as Guangzhou have already consulted foreign guidelines and AD books and have shown interest in learning how AD issues have been tackled in the West. It can be seen that, according to the sample, an overwhelming majority of those audio describing are volunteers, aged 20-50 and most of them have a university education background, although not related to Translation. Their professional work experience is also diverse: most of them have only been audio describing for a few years and always in Standard Chinese. They do not usually receive training on how to audio describe and almost all of them audio describe the same product: films.

During our research we have identified six issues that could be improved and that would contribute to ensuring AD a long life in China. First, the Chinese government should define and require the presence of media accessibility in more fields and even quantify it whenever possible. Wu and Xie (2015) also consider the national law to be insufficient and lacking in clarity. They also believe that one law mentioning “narrations” is far from enough and that legislation should also specify who is responsible for the provision of this service and how it is going to be made available.

Second, the copyright issue should be addressed through legislation so that persons with sight loss can have access to newer films without having to attend live sessions. In addition, for those unable to leave their homes, recorded CDs that can be borrowed from associations are a good way to give them more independence.

Third, as is already happening in Hong Kong thanks to the courses given by Dr. Leung, Chinese Translation faculties should start including media accessibility in their Audiovisual Translation courses. This would raise more social awareness and would prepare future audio describers better, since they would have a more solid base.

Fourth, if AD were considered a business, this would ensure continuity and provide those who have been trained with job prospects. An example can be found in Hong Kong, where Dr. Leung has implemented an AD business model and makes her customers pay for her services. According to her, this helps to create a market for professional audio describers and, therefore, ensure that there are professionals that offer this accessibility service.

Fifth, a Chinese official standard on how to create AD should be drafted. It should be based on empirical research that caters to users' needs and preferences. This would facilitate training sessions enormously and ensure homogeneity across the country. Leung's (2018) proposal could be a starting point but would also need to be tested in Mainland China.

Finally, since there are many different AD-related initiatives in Mainland China, Taiwan and Hong Kong, establishing some official channels of communication would also be advisable. A national conference on Chinese AD could be organized on a regular basis so that all stakeholders involved can exchange their views on the present situation and work together on a better future. All these recommendations should be added to those made by Wu and Xie (2015), who found that their informants thought the government should be the body responsible for extending the service in the country. These researchers also

suggested that the Chinese government should invest more money in the service, offer more professional training and develop more AD-related technology.

While this study clearly spells out the situation, it is not without its limitations. The size of our sample is not big enough to draw conclusions about the whole of Mainland China. It is also not a random sample: Shanghai audio describers have greater presence in the study because this is the city where the researchers are based and where they normally conduct their research, however there are some cities which offer AD that have not been included. The proportion of AD voicers is far from equal to that of AD scriptwriters in the sample, which obviously has an impact on the results. Nevertheless, it is hoped that this first attempt to investigate who is currently audio describing in Mainland China will engender further studies which contribute to both the creation of a new academic field in the country and an improved and more professional service. It is also hoped that our findings will be useful to better serve those in China with sight loss, raise social awareness of the needs of this collective, and train audio describers who can ensure a quality service based on academic research.

Acknowledgements

This paper has been funded by the EasyTV project (GA761999) and the Department of Translation and Interpreting and East Asian Studies of the Autonomous University of Barcelona. The authors are members of the TransMedia Catalonia research group (2017SGR113) and Irene Tor-Carroggio is currently enrolled in the PhD Programme in Translation and Intercultural Studies at the Autonomous University of Barcelona.

References

- CHMIEL, Agnieszka & Iwona Mazur. (2012) "AD Reception Research: Some Methodological Considerations." In: Perego, Elisa (ed.) 2012. *Emerging Topics in Translation: Audio Description*. Trieste: EUT, pp. 57-74.
- DI GIOVANNI, Elena. (2018) "Audio Description for Live Performances and Audience Participation." *JoSTrans: The Journal of Specialised Translation* 29, pp. 189-211.
- FENG, Jiyuan. (2018) "Koushu Yingxiang de Banquan Guizhi Tanjiu." [Study on the copyright of AD]. *Henan Keji Zhishi Chanquan* 3, pp. 36-38.
- FRESNO, Nazaret. (2014) *La (re)construcción de los personajes filmicos en la audiodescripción Efectos de la cantidad de información y de su segmentación en el*

- recuerdo de los receptores*. Barcelona: Universitat Autònoma de Barcelona. PhD thesis.
- FRYER, Louise. (2016) *An Introduction to Audio Description*. London: Routledge.
- GAMBIER, Yves. (2003) "Introduction." *The Translator* 9:2, pp. 171-189.
- GAMBIER, Yves & Haina Jin. (2018) "Audiovisual Translation in China: A Dialogue between Yves Gambier and Haina Jin." *Journal of Audiovisual Translation* 1:1, pp. 26-39.
- HERNÁNDEZ-BARTOLOMÉ, Gustavo & Ana Isabel Mendiluce-Cabrera. (2009) "How Can Images Be Translated? Audio Description, a Challenging Audiovisual and Social Gap-filler." *Hermeneus, Revista de Traducción e Interpretación* 11, pp. 1-17.
- JANKOWSKA, Anna. (2015) *Translating Audio Description Scripts: Translation as a New Strategy of Creating Audio Description*. Bern: Peter Lang.
- LEUNG, Dawning. (2018) *Audio description of audiovisual programmes for the visually impaired in Hong Kong*. London: University College London. PhD thesis.
- LEUNG, Dawning. (2015) "Audio Description in Hong Kong." In: Baños, Rocío & Jorge Díaz-Cintas (eds.) 2015. *Audiovisual Translation in a Global Context*. London: Palgrave Macmillan, pp. 266-281.
- LI, Dongxiao. (2013) *Tingjian, Kanjian: Yingshi Meiti de Wuzhang'ai Chuanbo Yanjiu* [Hear and See: Audio-Visual Media for All]. Hangzhou: Zhejiang Daxue Chubanshe.
- LI, Dongxiao & Xianghui Pan. (2013) "Huishenghuise: Zhongguo Wuzhang'ai Dianying de Fazhan Xianzhuang ji Zhanwang." [Vivid and realistic: Development and prospects of Chinese barrier-free movies]. *Zhejiang Xuekan* 4, pp. 188-198.
- LI, Dongxiao & Peter O. Looms. (2016) "Television Accessibility in China." In: Biswas, Pradipta; Carlos Duarte; Patrick Langdon & Luis Almeida (eds.) 2016. *A Multimodal End 2 End Approach to Access*. Cham: Springer, pp. 261-273.
- MATAMALA, Anna; María Machuca & Antonio Ríos. (2018) "More than Words. Voice in Audio Description." Paper presented at Languages and the Media Conference (October 2018), Berlin, Germany. Electronic version: <<https://ddd.uab.cat/record/196126>>
- ORERO, Pilar; Stephen Doherty; Jan-Louis Kruger; Anna Matamala; Jan Pedersen; Pablo Romero-Fresco; Sara Rovira-Esteva; Olga Soler-Vilageliu & Agnieszka Szarkowska. (2018) "Conducting Experimental Research in Audiovisual Translation (AVT): A Position Paper." *JoSTrans: The Journal of Specialised Translation* 30, pp. 105-126.
- PENG, Xiaoxia; Shige Song; Sheena Sullivan; Jingjun Qiu & Wei Wang. (2010) "Ageing, the Urban-Rural Gap and Disability Trends: 19 Years of Experience in China - 1987 to 2006." *PLoS ONE* 5:8, pp. e12129.
- PETERSEN, Carole J. (2008) "China's Ratification of the Convention on the Rights on Persons with Disabilities: The Implications for Hong Kong." *Hong Kong Law Journal* 38, pp. 611-643.
- QIU, Zhuoying. (1998) "Disability Statistics in the People's Republic of China." *Asia & Pacific Journal on Disability* 3.
- REVIERS, Nina. (2016) "Audio description services in Europe: an update." *The Journal of Specialised Translation* 26, pp. 232-247.
- SNYDER, Joel. (2014) *The Visual Made Verbal: A Comprehensive Training Manual and Guide to the History and Applications of Audio Description*. Arlington: American Council of the Blind. Abingdon: Routledge.
- VARIOUS AUTHORS (United Nations). (2006) *Convention on the Rights of Persons with Disabilities*. Electronic version: <<http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf>> .

- VARIOUS AUTHORS (China Disabled Persons' Federation). (2007) *2006 Nian Di-Er Ci Quanguo Canjiren Chouyang Diaocha Zhuyao Shuju Gongbao* [Report on the results of the second national sample survey on persons with disabilities]. Electronic version: <http://www.cdpf.org.cn/sjzx/cjrgk/200711/t20071121_387540_1.shtml>
- VARIOUS AUTHORS (China Disabled Persons' Federation). (2016) *Law on the Protection of Persons' with Disabilities*. Electronic version: <http://www.cdpf.org.cn/english/Resources/lawsregulations/201603/t20160303_542879.shtml>
- VARIOUS AUTHORS (People's Republic of China) (n.d.). *Copyright Law of the People's Republic of China*. Electronic version: <<https://www.wipo.int/edocs/lexdocs/laws/en/cn/cn001en.pdf>>
- VARIOUS AUTHORS (The National People's Congress of the People's Republic of China) (n.d.) *Beijing shi shishi "Zhonghua Renmin Gongheguo Canjiren Baozhangfa" banfa* [Beijing's implementation of the Law on the Protection of Persons with Disabilities]. Electronic version: <http://www.npc.gov.cn/npc/zfjc/cjrbz/fz/fjc/2012-07/26/content_1731040.htm>
- XU, Jing. (2018) *Zhongguo Mangwen Tushuguan Koushu Yingxiangguan Bangzhu Shizhang Renshi "Ting" Dianyong* [The AD screening room of the China Braille Library helps people "listen" to movies]. Electronic version: <<https://wemedia.ifeng.com/56470848/wemedia.shtml>>
- WALCZAK, Agnieszka. (2017) "Audio Description on Smartphones: Making Cinema Accessible for Visually Impaired Audiences". *Universal Access in the Information Society* 17:4, pp. 833-840.
- WU, Zongyi & Zhenzhen Xie. (2015) "Zhongguo Shizhang Koushu Yingxiang Fuwu de Fazhan Xianzhuang yu Dalu Tuiguang" [The development and the promotion of audio description in Mainland China]. *Xinwen Yanjiu Shikan* 6:10, pp. 20-22.
- YEUNG, Jessica. (2007) "Audio Description in the Chinese World". In: Díaz-Cintas, Jorge; Pilar Orero & Aline Remael (eds.) 2007. *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language*. Amsterdam: Rodopi, pp. 231-244.

**Chapter 3. Article 2: Chinese persons with sight loss'
opinion on their experience with audio description**

The customer is always right: Study on Chinese persons with sight loss' opinion on their experience with audio description¹⁶

Abstract

Audio description (AD) is in its infancy in China, both in terms of the provision of the service and as an academic object of study. This paper provides a useful insight into what Chinese AD users think of this access service and how they consume it. This information was gathered through two different questionnaires, mostly distributed in the city of Shanghai, although not exclusively. The results show what the profile of the AD user is, how satisfied they are with the service that is currently available and their expectations for the future. Only after identifying the areas that matter to users the most can both science and industry contribute to guaranteeing their rights in the best possible way.

Keywords: media accessibility; audio description; China; user satisfaction

3.1. Introduction

Confucianism, Buddhism and Taoism have played a key role in shaping the concept of disability in China. In early China physical abnormality was not inexorably identified with negativity, as popular deities often had uncanny body shapes (Zhou 2002, 105). Also, the *Zhuangzi* contains some passages that suggest that physical disability was not perceived as negative (Zhou 2002; Lewis 2014; Lambert 2016). Yet, this changed and throughout history Chinese persons with disabilities have been seen as not worthy of attention, as the result of karma (Campbell and Uren 2011) or their parents' defects (Palmer 2014), as a deficient fusion between man and nature (Avery 2016), as criminals, as outsiders and even as racial degeneration and one of the key causes for the nation's backwardness (Zhou 2002), just to name a few. It is, therefore, not surprising that persons with disabilities try to run away from traditional stigmatization. After interviewing persons with disabilities in China, Lin and Yang (2018) concluded that their interviewees shared a common feature: they denied being victims of any discrimination and, thus, did

¹⁶ Tor-Carroggio, I. (2020). The customer is always right: Study on Chinese persons with sight loss' opinion on their experience with audio description. *Disability & Society*. doi: 10.1080/09687599.2020.1713727

not pursue social inclusion. According to these researchers, this phenomenon is caused by a host of institutional, environmental, social, family-related and psychological factors, most of which have been inherited from previous generations' prejudices.

Like in any other language, the choice of words in Chinese reflects social attitudes towards certain issues, among which disability can be found. In China, the terms referring to it have negative connotations, although Aguado Díaz (1995, 277) observed that some social practices carried out in the past, like foot binding, evidence the cultural relativity of physical disability and its context-sensitiveness. Currently, the most widespread word to refer to persons with disabilities is *canjiren* (残疾人), which literally means “a person who is deficient and ill”. Other words that clearly exemplify how medicalized the conception of disability is in China are, for example, *canfei* (残废, “crippled”) and *shazi* (傻子, “simpleton”). Yet, new voices have emerged that advocate for new terms to refer to those functionally diverse, such as *canzhang renshi* (残障人士), “a deficient and handicapped person” (Hallett 2005; Palmer 2014). Unfortunately, although it sounds more formal, it still frames disability within the medical model of disability, which believes that disability stems from biological problems that must be addressed and that only concern individuals. This model has already been condemned by, for instance, the International Telecommunication Union (2017), which endorses an approach that vindicates the idea that the problem emanates from social behaviour, instead of the person with an impairment (Orero and Tor-Carroggio 2018).

The establishment of the China Disabled Persons' Federation by the Chinese central government in 1988 was the first step towards the development of policies and legislation on disability (Zhang 2017). Also, since China's ratification of the United Nations' Convention on the Rights of Persons with Disabilities in 2007, the country has made visible efforts to improve the life of persons with disabilities, such as the amendment of the Law on the Protection of Persons with Disabilities, originally passed in 1980. This is no surprise taking into account that “the occurrence of important domestic and international events often becomes a critical policy window for the improvement of the welfare of disabled people” (Tang and Cao 2018, 1173). Thus, the almost 83 million persons with disabilities in China are also benefiting from the country becoming more aware of their needs. One of the rights the United Nations' Convention on the Rights of Persons with Disabilities recognizes—hence the need of China to give an answer to it—

is the access to culture and information. In the case of persons with sight loss, who make up the group this paper revolves around, this right can be granted through sensory accessibility services such as audio description. Audio description, an exponentially-growing Audio-visual Translation mode, is a means of translating visual elements into aural description primarily for helping persons with sight loss to enable them to access live and pre-recorded programs and performances. In other words, it is the visual made verbal (Snyder 2014).

In fact, persons with no sensory loss receive information from two sensory modalities: sight and hearing. Audio description interacting with the original sound track “should stimulate a number of other senses in the user” (Fryer 2016, 47) and therefore grant users a similar experience to that of sighted viewers. Some users have even argued that their experience is enhanced because the lack of visual reminders; such as a TV screen; eradicates the need for suspension of disbelief (Fryer 2016, 47). This accessibility service is, along with other such as subtitles for the deaf and hard of hearing and sign language interpreting, already legally required in countries such as Spain, the UK, and even Japan—although to a different extent (Martínez-Sirés 2016, 34).

In contrast, audio description is still at its infancy in China, both as a professional service and as a research topic (Tor-Carroggio and Casas-Tost, forthcoming). Martínez-Sirés (2016, 27) had also highlighted the lack of scholarship in media accessibility—an umbrella label for a number of sensory access services—for non-Western languages. Known as *koushu yingxiang* (口述影像, “image description”) and *wuzhang'ai dianying* (无障碍电影, “barrier-free movies”), depending on the province and on the level of accuracy aimed at, some researchers have tried to contextualize audio description from a historical point of view (Li 2013) and even by reporting its state of the art and profiling the persons in charge of delivering it in China (Tor-Carroggio and Casas-Tost, forthcoming). Nonetheless, one of the main angles from which audio description needs to be studied remains unveiled: user needs and satisfaction. In fact, Matamala and Orero (2018) consider this to be of the utmost importance and underline the need of making the end user's voice heard, and that research supports the statements they make.

This paper aims at bridging this existing research gap by presenting the results of two different surveys conducted in Mainland China in 2018. Both questionnaires departed from the following three research questions:

- 1) What are the habits of audio description users when it comes to this accessibility service?
- 2) What are the needs of audio description users when it comes to this accessibility service?
- 3) How satisfied are users towards the audio description currently delivered?

This article is divided into four sections. The first one revises previous studies that have focused on Chinese persons with sight loss and Chinese media accessibility. The second one explains the methodology followed in the two studies presented in this paper. The third section analyses the results of the two surveys conducted with audio description users. Finally, the discussion and conclusions are put forward.

3.2. Research background

Although the events related to it are reported in the press from time to time, media accessibility is still an under-researched field in China (Gambier and Jin 2018). This is also proved by doing a quick search combining “Chinese” and “media accessibility” in the Translation Studies database BITRA, which is the field, more specifically Audiovisual Translation, in which media accessibility is encompassed. Currently, such a search only yields four results. Something similar happens when consulting the Chinese academic database CNKI. Consequently, it is not surprising that only a few scholars have investigated audio description users to find out what their needs are and how audio description should be delivered to suit them best. Chao (2002) and Liu (2015) carried out some research in Taiwan in this regard. The former was the first academic researcher to consider audio description as a research topic in Greater China, while the latter studied the need of using audio description to assist the learning of elementary school students with visual impairment. This is also the case of Leung (2018), who examined the media use, behaviour, and motivations—as well as the reception and preferences of the visually impaired audiences when consuming audio description in Hong Kong. Similarly, Li (2013), in a rather comprehensive study on media accessibility in Mainland China, included a section presenting the results of a survey administered to persons with sight loss (N = 729) in Zhejiang province from which their media use, behaviour, and satisfaction were elicited. Her research confirmed that the most popular means of communication amongst them were television (86.6%), radio (75.3%), mobile phones

(68.1%), the Internet (45.3%), DVD/VCD (24.8%), movies (21.2%), and newspapers (20.5%). The respondents claimed to use the media primarily to get informed (64.0%) and to entertain themselves (41.3%), and they were “quite satisfied” (51.6%) with how it met their needs. As for their habits regarding the TV, 50.0% of the informants watched it more than one hour every day and the most popular type of program was the news. Li connected the latter finding with the fact that news programs allow persons with sight loss to be less dependent on the images being broadcast. The fact that audio description is not incorporated in TV programs did not prevent Li’s (2013) respondents from affirming they were satisfied with TV content (Li 2013, 59-60). As for the lack of audio description on TV, Li and Looms (2016) investigated the current incompatibility between Chinese TV sets and audio description:

One of the reasons for not including the capability to offer closed captions or audio description in China’s digital television standard, China Multimedia Mobile Broadcasting (CMMB) 10 years ago was a concern about the affordability of set-top boxes (STBs) [sic]. Industry sources suggest that the retail price of STBs had to be kept below RMB 200 (currently USD 30). As the inclusion of closed captioning and audio description would have required patent licenses, this would have driven costs above the target retail price (Li and Looms 2016, 267-268).

As for movies, which are the only field in which audio description is applied in China, Li’s (2013, 41-47) results show that they are a less popular choice amongst blind and partially-sighted persons. More than half of the respondents (54.0%) never watched movies mainly because of the inconvenience of getting to the venue.

The scarcity of academic bibliography might derive from society’s general lack of interest in this group of persons, as the results from Wu and Xie’s (2015) study revealed. These researchers conducted an online survey in China and found out that, out of the 160 valid responses gathered, 79.38% of the respondents admitted not caring about persons with sight loss. This is probably why 88.75% of them did not know about the existence of audio description (Wu and Xie, 2015). The lack of literature can also be attributed to persons with sight loss not having attached enough importance to or not having vindicated enough their right to culture and information because of them having to deal with other difficulties. In fact, Li’s (2013) survey proved that the main challenges respondents had to face were poverty (24.7%) and loneliness (24.7%). She even recounted in her book how users questioned her research for not focusing on what they considered more critical problems, such as employment. Also, the scant research in this field has not been compensated from abroad, as, as far as we know, only Fei (2011) wrote a masters

dissertation on audio description and the Chinese community in Canada, which aimed at examining the differences and similarities of expectations and evaluations of integrated audio description between the Chinese and Canadian communities. Yet, his results need to be contextualized, since the stimuli he used was audio description in English with both groups.

It is in this context of lack of studies on audio description in China that this research is framed in, with the aim of making a step forward from the users' perspective.

3.3. Methodology

This research presents the results of two different questionnaires that catered for the same target respondents – Chinese persons with sight loss that had experience with audio description.

The first questionnaire was inspired on an existing one created by the European project ADLAB PRO (<http://www.adlabproject.eu/>), which gauged user satisfaction and opinions on audio description. It was “inspired” and not “based” on that survey because numerous changes had to be made to fit the Chinese case. For example, “opera” was replaced for “Chinese opera” in those questions in which some fields were listed so that users chose where they wanted to enjoy audio description. The questionnaire had also to be shortened because the circumstances in which the questionnaire was expected to be delivered would not allow us to enter into much detail. The questionnaire was then orally piloted in a blind massage parlour on 18 April 2018 in Shanghai. Five blind masseurs were recruited and only one had previous experience with audio description. Yet, all their comments provided useful insight into the design of questionnaires for this respondent profile. The pilot resulted in many changes, such as adding a clearer explanation of what audio description was in the information sheet and making all the information and questions more easy-to-read. This was attributed to the users' low educational level and lack of knowledge about this accessibility service. It was also found that these users did not know the Chinese term for audio description and an explanatory remark was made to show that the access service delivered by the so-called “barrier-free movies” was also known as *koushu yingxiang*. Also, it was confirmed that users had difficulty in understanding 1-10 scales and were consequently substituted by categorical ones. All these changes arising from the pilot were considered to be a normal practice, as, for

instance, Chmiel and Mazur (2012) indicated they had to change their testing materials up to sixteen times before formally administering them. Before conducting our survey, and since applying for ethical approval is becoming commonplace in the execution of Audio-visual Translation research (Orero et al. 2018), our questionnaire was approved by our university's Ethics Committee. Ethical considerations in human research have been a major concern since the critical articles written by Pappworth (1967) on medical research. The final questionnaire was administered at the Cathay Cinema in Shanghai on 26 April 2018. This cinema is one of the sixteen cinemas that offer a live audio description session for free thanks to the financial help of the local government (Tor-Carroggio and Casas-Tost, forthcoming). With the help of some volunteers, who had been briefed on what audio description was and also on the aim of the research, cinema-goers with sight loss were approached while they were lining up to get into the cinema for their monthly live audio description session. This was thought to be the best way to find as many users as possible, since one of the requirements to answer the questionnaire was to be or to have been an audio description user. An online version of the questionnaire was also made available in the website Web Survey Creator. It was sent to the China Braille Library in Beijing so that they helped us recruit informants from a different city. Therefore, the survey was administered orally and through the Internet.

The second questionnaire was both designed and administered by the association Sound of Light Barrier-free Film & TV Culture Development Centre (abbreviated from now on as "Sound of Light") in Shanghai. Sound of Light is an association led by Han Ying which was set up in Shanghai in 2016. It receives direct material and financial assistance from the Shanghai Association of Persons with Disabilities. They produce about 50 audio described movies every year, which are recorded in their headquarters and distributed through digital means. The movies are audio described thanks to a network of volunteers that both write the audio description script and voice it. These movies can be accessed at specific projecting sites which are all located in Shanghai (Tor-Carroggio and Casas-Tost, forthcoming). The aim of the questionnaire was to gather data about the experience, if any, of persons with sight loss regarding the so-called "barrier-free movies", as well as their opinion on them. In this case, the survey was only conducted in Shanghai. It was administered in person and orally by volunteers and by the association's staff before or after AD sessions. The surveys were collected by Sound of Light employees and stored

in the association's headquarters. Data collection started in April 2018 and was extended until January 2019.

The reasons for presenting the results of both surveys in the same paper are the following. First, the first questionnaire's sample was smaller than the second one, so it was thought that the first findings could be contrasted and complemented with the data obtained from the second survey. Yet, the first questionnaire could not be replaced by the second one because the questions asked were not the same and both provided significant information about user experience and satisfaction towards audio description. Second, despite the fact that the second questionnaire was not designed by myself, I was offered the data to analyse it in order to ultimately present the results to the Shanghai Association of Persons with Disabilities. Finally, as it has already been underlined, this is the first time this kind of data from Mainland China is being presented to academia and the aim of the paper was to compile what is currently available and offer as complete of a picture as possible of what has been done so far in this regard.

For the statistical analysis IBM SPSS (version 22) was used.

3.4. Results of survey 1

The first questionnaire that was administered was framed within PhD research aimed at verifying if Chinese persons with sight loss would accept the use of artificial voices in the audio description of films. Before doing so, audio description users had to be studied and profiled to better design the next stages of the PhD research.

3.4.1. Informants' demographic profile

Fifty-two informants answered this first questionnaire. They were from both Shanghai (42) and Beijing (10). The average age was 63.8 years old and the median was 65 years old. The youngest informant was 28, while the eldest was 88. It is worth highlighting that 17 persons decided not to answer this question. In the light of the average age observed, it was not surprising that most of our respondents were retired. Almost 6% of our respondents were unemployed, 1.9% were freelancers and 3,8% of the persons who answered our questionnaire stated that they had a different employment situation than those listed.

Figure 3.1. clearly shows that only a minority (3.8%) had attended university, with secondary education being the educational level which the majority of informants had completed.

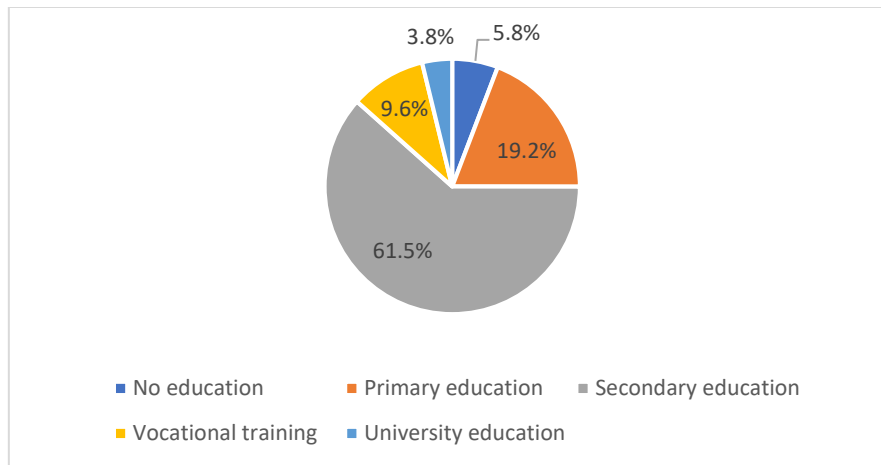


Figure 3.1. Informants' educational background

As regards their mother tongue, more than half of the respondents (57.7%) claimed to speak a Chinese dialect, although Standard Chinese (42.3%) is the language in which movies are mainly audio described in China. The average level of Standard Chinese was found to be 0.83 out of 2. A reason that could explain our finding could be the well-known Chinese modesty and cautiousness. Yet, Rovira-Esteva (2010, 257) also noted that, according to official data from 2004, the older Chinese persons are, the less they can use Standard Chinese to communicate. The same happens with their academic background: the more educated they are, the better they can communicate in the country's lingua franca.

With respect to their visual condition, 67.3% of the informants claimed that their sight loss was congenial, whilst 23.1% said it was acquired.

3.4.2. Informants' experience and opinion on audio description

All the informants were consumers of audio described films, and more than 80% did so in cinemas, but none of them selected any of the other areas in which audio description can be provided, i.e., TV, museums, among others. Yet, they would be interested in having the service available in different fields, as Figure 3.2. points out.

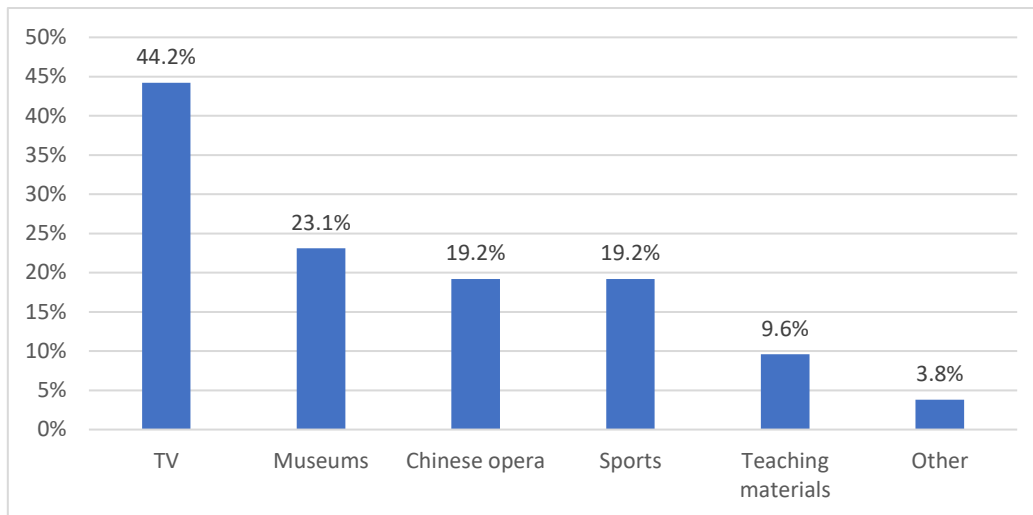


Figure 3.2. Fields in which respondents would like to have AD

TV revealed itself as the most sought-after mode for audio description, although the distribution obtained was not statistically significant ($\chi^2 = 0.692$, $p = 0.405$). This meant that a distribution in which “yes” and “no” had the same weight could not be discarded.

As for the frequency of use of the service, three quarters of the informants claimed to consume audio description in films frequently or very frequently. 15.4% did so sometimes and only 9.6% of the informants declared that they rarely watch audio described films.

One of the main points of the questionnaire was to determine the degree of user satisfaction with audio description. It was observed that our informants were more satisfied with the quality of AD (mean = 3.33, in a 0-4 scale) rather than with its quantity (mean = 3.15, in a 0-4 scale). In order to determine if satisfaction in terms of quantity was statistically different to that in terms of quality, a Wilcoxon signed ranks test was carried out. In spite of the means showing some slight differences, it is not possible to state that users are more satisfied with audio description quality than with the quantity offered ($Z = -1.415$, $p = 0.157$, with 33 ties).

It was also deemed interesting to see if there was any correlation between the frequency of audio description consumption and user satisfaction, both in terms of quality and quantity. A statistically significant linear relation was found –although weak– between audio description consumption frequency and user satisfaction with quantity ($r_s = 0.437$, $p < 0.05$).

The last question sought the opinion of users about which types of movies they considered more suitable to carry out some experiments with artificial voices. Although this

experiment is not the focus of this paper, the results of this question are also added because they can be understood as the genres users liked the most. Respondents could select more than one option. The top three choices were historical (36.5%), war (25.0%), and inspirational (25.0%) movies (Figure 3.3.).

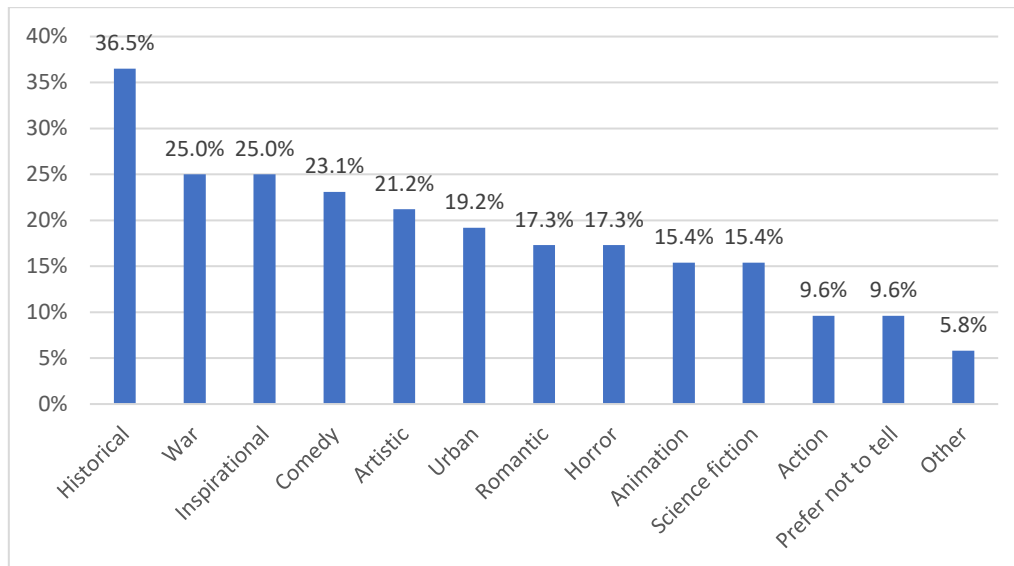


Figure 3.3. Informants' preferred movie genres

3.5. Results of survey 2

The first section presents their demographic profile, whereas the second one focuses on their experience and opinion on audio described movies.

3.5.1. Informants' demographic profile

One hundred and fifty-six persons answered this questionnaire. The proportion of women (56.4%) was slightly higher than that of men (41.7%). Just about half of the informants (50.6%) were 60 years old or above. Less than 1% were 18-25 years old, those aged 18-35 only represented 4.5% of the respondents, the persons aged 36-45 represented 12.8% of the respondents and almost a third were 46-59 years old (30.1%). Three persons did not answer this question.

As for the informants' academic background, the vast majority did not attend university (92.9%), as Figure 3.4. shows. This is in line with the age of most informants, who probably did not have a chance to access even basic education in their youth because of their physical impairment or the socio-political situation in the country at that time.

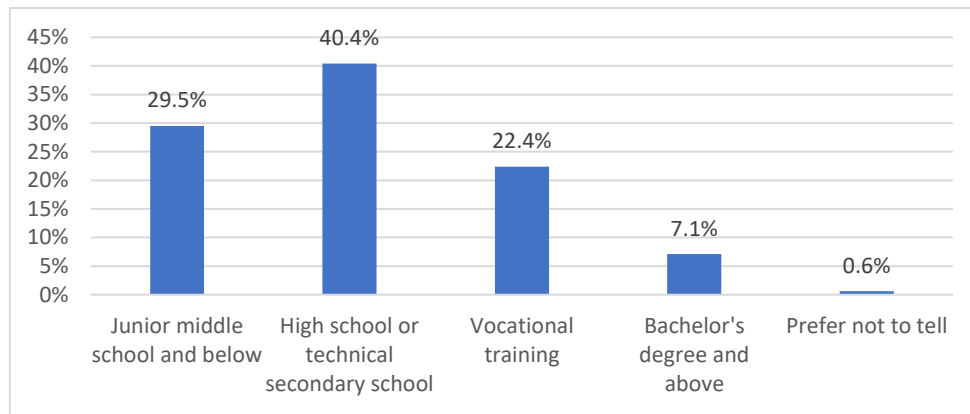


Figure 3.4. Informants' education background

Regarding their employment situation, it can be stated that there were no students among the informants. More than half of the respondents were retired (66.7%). The rest were either employed or depended on their families (28.2%) or had other situations (4.5%).

3.5.2. Informants' experience and opinion on audio description

Regarding whether informants were aware of the existence of audio described movies, 90.4% of the respondents claimed to know about it, whereas 12 persons (7.7%) admitted they did not know and three persons refused to answer this question. This result makes sense when taking into account the fact that those surveyed were recruited during activities related to the so-called "barrier-free movies".

When asked about how they came to know about this service, more than 80% revealed that they had learned about it through user associations. The rest of the informants mentioned "assistive staff" (23%), "family and friends" (16%), and "the media and the Internet" (15%). 20 persons did not answer this question.

Regarding whether respondents had the need of audio described movies, an overwhelming majority thought so (94.2%) and only three persons preferred not to answer. Those who answered "no" in this question were asked to stop filling in the survey.

The remaining 148 informants were then asked what their purpose was when watching audio described movies. They were allowed to select more than one option. The choices given were "for entertainment" (67.3%), "to socialize" (37.8%), and "for study purposes" (49.4%).

Informants were also asked if they had ever watched movies with audio description. Almost 84% of the respondents had watched them (83.3%). These persons were mostly satisfied or very satisfied (76.3%) and only one person admitted not being satisfied with them.

The questionnaire was also interested in knowing how respondents got to the place where barrier-free movies are screened. More than half of the 128 respondents that answered claimed to use public transport, such as the bus or the metro (64.1%). Less than a quarter of respondents (17.3%) stated they went on foot and one informant chose the option of "others".

When asked whether there were parts of the audio described movies that were not understood while watching them, more than half of the respondents said that there were almost no such parts (55.1%). Some respondents admitted that they sometimes had to face this situation (26.9%) and only one person admitted that this situation happened quite frequently. Twenty-seven informants preferred not to answer.

To the question about how users wanted the information regarding barrier-free movies to be delivered to them. The option of reaching out to them through Wechat or QQ, both instant chat tools, was the most popular one (30.8%) among the 125 informants that answered the question, closely followed by a Wechat public account (26.9%). Door-to-door notification was only chosen by 16.7% of the informants and nine persons selected the option of "others".

Informants were asked what kind of movies they liked the most. They were able to choose three of the 12 types presented. The top five were comedies (47.4%), historic films (37.8%), inspirational films (33.3%), romantic films (27.6%), and war films (25.6%), as Figure 3.5. shows.

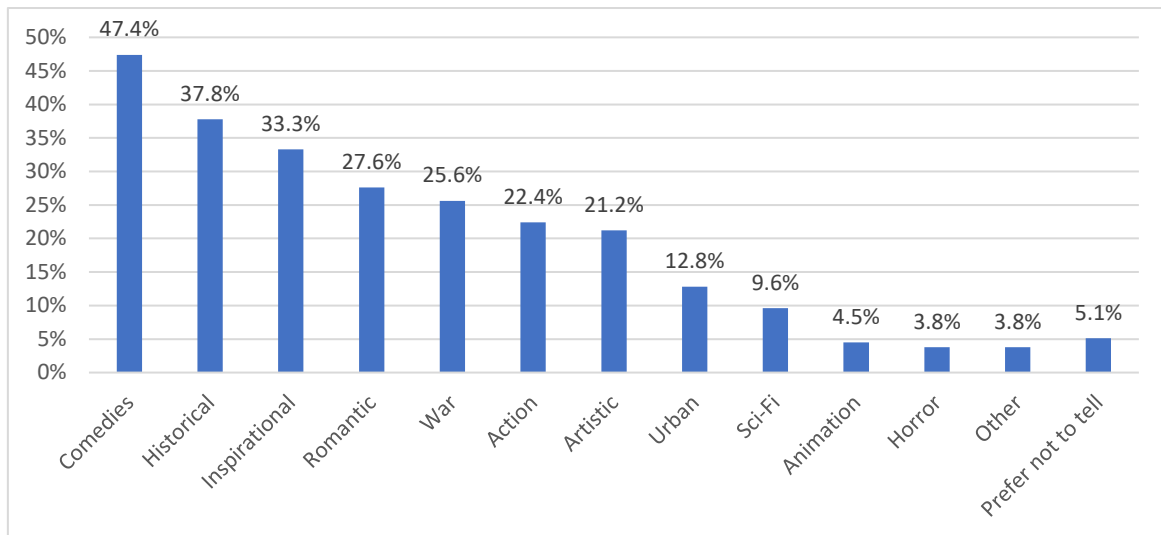


Figure 3.5. Informants' preferred movies

Participants were also asked what type of movies they preferred: classic or new. The option of “both” was also available. Almost 40% of the respondents (39.7%) said that they liked classic movies the most, whereas only 14.1% stated they enjoyed newer movies more. Still, 37.2% of respondents affirmed they liked both types. 14 persons refused to answer.

Regarding this last question, an association between the age and the type of movie chosen was found to be statistically significant ($\chi^2 = 20.39$, $p = 0.009$). This means it can be affirmed that, in general, the older a person is the more they prefer to watch old films. In our case, it was observed that persons who were 60 or older preferred old movies.

Finally, a qualitative question was added to allow informants to make any remarks they thought were appropriate. First, some respondents thought that audio described movies needed to be screened more often and better publicized so that more users could benefit from them. Second, it was considered necessary to turn down the volume of the original soundtrack because it sometimes did not allow users to hear the audio description clearly. Third, a few informants wanted the types of movies screened to be more varied and to include more foreign films. Fourth, some persons also asked for more staff to help users when entering and leaving the facilities. Finally, many users hoped that in the future persons with and without sight loss could watch movies together at the cinema. They thought this could be achieved by persons with sight loss using headphones to listen to the audio description.

3.6. Discussion

These two surveys have helped us draw a picture of the users' profile of audio description in China. Although in both cases the sample was quite aged, being a limitation of this study, the users' similar profile allowed us to draw justifiable comparisons among the two questionnaires. Also, despite the sample undoubtedly not being representative of the whole China, it must be stated that vision problems are mostly age-related (Fryer 2016, 43), so the samples, even though they can be criticized, can also shed some light on the majoritarian group of audio description consumers. It could be deduced from the average age in both questionnaires that audio description is mostly consumed by elderly persons, but the truth is that audio description sessions for children are also carried out and recorded material is produced for them as well. Fryer (2016, 43) also claims that women are more likely than men to become blind simply because they live longer. This is supported by the slightly higher proportion of women in the sample recruited by Sound of Light. The first questionnaire did not ask about informants' sex because it was not deemed necessary to assess user experience. The demographic similarities between the two samples go beyond age, since in both cases more than 90% of the informants did not attend university and the clear majority were retired.

All the informants were consumers of audio described movies and none mentioned any kind of experience with other types of audio description. Leung's (2018) results also point out that audio description is mostly consumed in films in Hong Kong, although it has a wider scope in the region and, therefore, can also be found in theatres and art exhibitions, for example. All our informants watched audio described movies quite frequently, which can be explained by the fact that their age allows them to have a lot of free time. In contrast, Leung's (2018) informants did so mostly sometimes (38.6%) or rarely (27.3%). Thanks to the Sound of Light questionnaire we found out that users perceived audio description mainly as an entertainment tool. This is in line with Leung's (2018) informants, who also claimed to watch films to entertain themselves. Yet, the respondents of the Sound of Light questionnaire also revealed having other objectives when watching audio described movies, such as socializing and learning, which is in line with studies carried out in other countries, such as that of Fryer (2016, 49), who had already pointed out that users valued audio description as a means of social inclusion. In fact, Li (2013) had already found out that one of the problems persons with sight loss had to face the most was loneliness, along with others such as poverty. Macdonald et al. (2018) framed persons with disabilities'

loneliness in the social model of disability and argued that so many of them feeling that way was due to the presence of certain disabling barriers, not because of their impairment.

As for user satisfaction with the audio description available so far, it seems that they think quite highly of it, which might seem strange taking into account that mainly volunteers with no professional training deliver it (Tor-Carroggio and Casas-Tost, forthcoming). This can be understood in different ways. Since persons with disabilities have traditionally been ignored and not taken care of in China, it is possible that any initiative that aims at helping or including them is highly praised, no matter the level of professionalism with which it is carried out. Yet, Branje and Fels (2012) found out that amateur description could maintain an acceptable level of quality so the possibility that the audio description prepared in China being of good quality; despite its conditions; is also more than feasible.

Last but not least, users preferring historical films in general was probably related to their age, since the top ten highest-grossing 2018 films in the Chinese market were consulted and none belonged to that genre. Similarly, Leung (2018) found out that historical movies were the second preferred genre in her study, in which 44 persons with sight loss, mostly aged above 40, participated. Therefore, in case this information is used for the preparation of future experiments or even for the real delivery of the service, it must be taken into account that the movie genre probably cannot be divorced from the age of participants.

3.6.1. The European case

Since the first questionnaire presented was originally used in the ADLAB PRO project, it is possible to compare our results with some of those obtained in the European case. In spite of the respondents coming mostly from 23 European countries, some non-European respondents coming from, for example, Brazil and Mexico were recruited as well (ADLAB PRO 2017). Although the European results do not represent the whole continent, the picture is still quite valuable because the most represented countries in the survey (mainly the UK, Spain and Italy) are those in which audio description has already come of age.

According to ADLAB PRO (2017), the original questionnaire was drawn up in all the project languages, i.e. English, Italian, Polish, Spanish, Slovene and Dutch. One hundred AD users responded the questionnaire but, unlike the respondents from our sample, not

all of them were persons with sight loss (5%). This detail also illustrates the fact that in Europe it is already common to believe that, although mainly targeted at persons with sight loss, audio description can also benefit sighted viewers. This is even endorsed by the International Organization for Standardization's Technical Specification 20071-21 "Information Technology — User interface component accessibility — Part 21: Guidance on audio description" (ISO 2014).

The sample recruited was more representative than the ones presented in this paper in terms of age, since they managed to find a more balanced number of respondents for each age bracket. Also, it is worth mentioning that 51% of the respondents had attended university. This stands in stark contrast with the educational background of our Chinese informants, almost none of whom had attended university. Had the Chinese samples included more young respondents, this could have been different, since Tang and Cao (2018, 1172) state that after cases that hit the headlines such as that of Dong Lina, education institutions are committed to provide special arrangements for persons with disabilities when they participate in any national examinations according to the newest revision of the People with Disabilities Education Ordinance.

Unlike our informants, the ADLAB PRO respondents admitted having access to more kinds of audio description, although most use it to watch films or TV, or when visiting a museum. AD is not used that often in the case of theatre, opera and other live events. Some Slovenian respondents complained about the inconvenience of time programmes with audio description are broadcast or even the lack of it in theatres or opera. However, it can be observed that China is lagging behind because audio description is still not even offered on TV, which is what users demand the most and is probably the most popular medium for audio description in many countries such as Poland (Jankowska and Walczak 2019).

Regarding end-user satisfaction, the European results showed that most users were dissatisfied (46%) or very dissatisfied (22%) with the quantity of audio description provided in their country, although they were quite satisfied with the quality of the products that were offered (14% were very satisfied and 58% were satisfied). We also observed users being more satisfied in terms of quality, but they also claimed to be quite satisfied with the quantity provided, probably because they are still not aware of all the applications audio description can have.

The original ADLAB PRO questionnaire inquired about the aspects related to audio description that users appreciated and disliked the most in regards textual and technical aspects, as well as AD language and style. Yet, due to the conditions under which the first questionnaire circulated, this kind of questions were left out. Focus groups or one-to-one interviews could be arranged in the future to go into these areas more in depth.

3.7. Conclusions

Chinese persons with disabilities rarely raise their voices to express themselves or to get engaged in rights advocacy (Zhang 2017). Yet, there are already some accessibility services mostly targeted at persons with hearing and sight loss with which access to information and culture, a right recognized in the United Nations' Convention on the Rights of Persons with Disabilities, can be granted. audio description is one of these services and in order to enhance its provision and academic research, which are rather scarce in China so far, users need to be involved. Not much is known about Chinese persons with sight loss, but these are key to developing audio description and making sure the service meets their needs.

This paper is the first of its kind, since it provides a deep insight into what Chinese audio description users think of the service and how they consume it. Only after identifying the areas that matter to users the most can science and industry contribute to guaranteeing their rights in the best possible way. Some conclusions can be drawn from the results of the two questionnaires carried out. First, old persons with little or no education and who are already retired seem to be the users that most attend audio description sessions in the cities where the sample was taken, Shanghai and Beijing. Second, they only have access to audio description in films, which they consume quite frequently mainly for entertainment purposes and thanks to the information provided by user associations. Third, they are all satisfied with the service provided so far, despite it not being delivered by persons who have been professionally trained in this field. Fourth, their favorite movie genres seem to be historical films and comedies.

Finally, it is worth highlighting that the results of the Sound of Light questionnaire had immediate consequences that had an impact on the audio description policy of the association. For example, in light of the fact that most of their users are elderly persons with little education, they came up to the conclusion that the terminology in the scripts

needed to be simplified. This could actually be a future line of research to explore, since, to the best of our knowledge, easy reading and audio description have never been combined before, although there are some EU-funded projects that have already started looking into that direction, such as the EASIT project (<http://pagines.uab.cat/easit/en>). Also, they decided to take into account the movie genres preferred when selecting the movie to be audio described.

Our study is not exempt of limitations. First, the first questionnaire was delivered in rather difficult circumstances, in which other cinema-goers would be expressing their opinion while one was being surveyed. This may have prevented the informant from expressing their true views by exposing them to peer pressure. Second, since Chinese volunteers helped with the administration of the first questionnaire, and although they were briefed on our research, one cannot be sure that all the questions raised by the informants were answered appropriately. Third, in both cases, especially the first one, the sample was rather small and was not obtained randomly but by recruiting persons attending the screenings of audio described films. Also, the respondents in both cases were mostly elderly persons coming mainly from one Chinese city. This causes us to not be able to draw conclusions for the whole of China. Therefore, future studies should include bigger and aleatory samples, integrated by persons coming from a wider range of places in China.

Although Chinese persons with sight loss have little experience with audio description, it is worth carrying out research to elicit their opinions on it, as Chmiel and Mazur (2012) underlined when mentioning countries in which audio description still had a lengthy path ahead of it. Also, working within the United Nations' Convention on the Rights of Persons with Disabilities paradigm means to consult end users, following the "nothing about us without us" approach. That is the reason why, in case audio description wants to be enhanced in Mainland China, the academic research concerning it must frame itself to a user-centric framework.

It is hoped that this paper will raise awareness on how important it is to involve persons with sight loss in the provision or research of anything that concerns them directly. Li and Looms (2016) assured removing the barriers persons with sight and hearing loss face clearly has a high priority in China, so it is expected that studies like this will set the course for future research because the need for it is overriding.

Acknowledgements

This article has been funded by the EasyTV project (GA: 761999) and the Confucius Institute. The author is member of the TransMedia Catalonia research group (2017SGR113) and is currently enrolled in the PhD Programme in Translation and Intercultural Studies at the Autonomous University of Barcelona.

References

ADLAB PRO. 2017. "Audio description professional: Profile definition". Accessed 2 September 2019. <https://www.adlabpro.eu/wp-content/uploads/2018/04/IO2-REPORT-Final.pdf>

Aguado Díaz, Antonio León. 1995. *Historia de las deficiencias*. Madrid: Escuela Libre Editorial, Fundación ONCE, 40–41.

Avery, Kevin. 2016. "Disability and the Three Traditional Chinese Belief Systems". Accessed 7 November 2018. <https://www.chinasource.org/resource-library/articles/disability-and-the-three-traditional-chinese-belief-systems>

Branje, Carmen, and Deborah Fels. 2012. "LiveDescribe: Can Amateur Describers Create High-Quality Audio Description?". *Journal of Visual Impairment and Blindness* 106(3), 154–165.

Campbell, Anne, and Marie Uren. 2011. "The Invisibles"... Disability in China in the 21st Century. *International Journal of Special Education* 26(1), 12–24.

Chao, Ya-li. 2002. 语言世界中的流动光影。口述影像的理论建构 [The moving lights in the world of language. Theory of audio description] (Yuyan shijie zhong de liudong guangying. Koushu yingxiang de lilun jiangou). Taipei: Wunan Publishing Co.

Chmiel, Agnieszka, and Iwona Mazur. 2012. AD Reception Research: Some Methodological Considerations. In *Emerging topics in translation: audio description*, edited by Elisa Perego, 57–74. Trieste: EUT.

Fei, Xue. 2011. *Audio Description in the Chinese Community*. Masters diss. Capital University of Economics and Business, Canada.

Fryer, Louise. 2016. *An Introduction to Audio Description*. London: Routledge.

Gambier, Yves, and Haina Jin. 2018. "Audiovisual Translation in China: A Dialogue between Yves Gambier and Haina Jin". *Journal of Audiovisual Translation* 1(1): 26–39.

Hallett, Stephen. 2005. "One Eye on China: Challenging Disability". Accessed 2 November 2018. <http://www.bbc.co.uk/ouch/features/one-eye-on-china-challenging.shtml>

ISO (International Organization for Standardization). 2014. "TS 20071-21 'Information Technology — User interface component accessibility — Part 21: Guidance on audio description'".

International Telecommunication Union. 2017. "Question 7/1: Access to Telecommunication/ICT Services by Persons with Disabilities and with Specific Needs". Accessed 12 October 2018. https://www.itu.int/dms_pub/itu-d/opb/stg/D-STG-SG01.07.4-2017-PDF-E.pdf

Jankowska, Anna, and Agnieszka Walczak. 2019. "Audio description for films in Poland: history, present state and future prospects". *JoSTrans: The Journal of Specialised Translation* 32: 236–261.

Lambert, Andrew. 2016. "The Usefulness of the Useless': Rethinking Disability through Classical Daoist Thought". In *Disability and World Religions: An Introduction*, edited by Darla Y. Schumm and Michael Stoltzfus. Waco: Baylor University Press.

Leung, Hoi Ching Dawning. 2018. "Audio Description of Audiovisual Programmes for the Visually Impaired in Hong Kong". PhD diss., University College London.

Lewis, Crismon. 2014. "Use of the Useless": Assessing Depictions of Disability in the Zhuangzi. Bachelor's diss. Brigham Young University. Accessed 20 November 2018.

https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1017&context=libraryrg_studentpub

Li, Dongxiao. 2013. 听见, 看见: 影视媒体的无障碍传播研究 [Hear and See: Audiovisual Media for All]. Hangzhou: Zhejiang University Press.

Li, Dongxiao and Peter Looms. 2016. "Television Accessibility in China". In *A Multimodel end 2 end approach to access*, edited by Pradipta Biswas, Carlos Duarte, Patrick Langdon and Luis Almeida, 261-273. Springer.

Lin, Zhongxuan, and Liu Yang. 2018. "Denial of Disability Discrimination by Disabled People in China". *Disability & Society* 33(5), 804–809.

Liu, Chih-Ching. 2015. 你是我的眼。運用口述影像輔助國小視障生學習之探究 [You Are My Eyes: Using Audio Description to Assist the Learning of Elementary School Students with Visual Impairment]". PhD diss., National Taiwan Normal University.

Macdonald, Stephen J., Lesley Deacon, Jackie Nixon, Abisope Akintola, Anna Gillingham, Jacqueline Kent, Gillian Ellis, et al. 2018. "'The Invisible Enemy': Disability, Loneliness and Isolation". *Disability & Society* 33(7): 1138–1159.

Martínez-Sirés, Paula. 2016. "Making the Inaccessible Accessible: A Panorama of Audiovisual Translation, Accessibility and Audio Description in Japan". *Asiadémica* 7, 24–50.

Matamala, Anna, Pilar Orero, Sara Rovira-Esteva, Helena Casas-Tost, Fernando Morales-Morante, Olga Soler-Vilageliu, Belén Agulló, et al. (2018). "User-centric Approaches in Access Services Evaluation: Profiling the End User". In *Proceedings of the Eleventh International Conference on Language Resources Evaluation (LREC 2018)*, edited by Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, H el ene Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, Takenobu Tokunaga, 1–7.

Matamala, Anna, and Pilar Orero. 2018. "Standardising accessibility: Transferring Knowledge to Society". *Journal of Audiovisual Translation* 1(1), 139–154.

Na, Tang, and Yu Cao. 2018. "From multiple barriers to a co-prosperity society: the development of a legal system for disabled people in China". *Disability & Society* 33:7, 1170-1174.

Orero, Pilar; Stephen Doherty; Jan-Louis Kruger, Anna Matamala, Jan Pedersen, Pablo Romero-Fresco, Sara Rovira-Esteva, et al. 2018. "Conducting Experimental Research in Audiovisual Translation (AVT): A Position Paper." *JoSTrans: The Journal of Specialised Translation* 30, 105–126.

Orero, Pilar, and Irene Tor-Carroggio. 2018. "User Requirements When Designing Learning e-Content: Interaction for All". In *Designing for the User Experience in Learning Systems*, edited by Evangelos Kapros and Maria Koutsombogera, 105–122. Cham: Springer Nature.

Palmer, James. 2014. "Crippling Injustice". Accessed 14 November 2018. <https://aeon.co/essays/what-is-life-like-for-disabled-people-in-china>

Pappworth, Maurice Henry. 1967. *Human Guinea Pigs: Experimentation on Man*. London: Routledge and Kegan Paul.

Rovira-Esteva, Sara. (2010). *Lengua y escritura chinas: Mitos y realidades*. Barcelona: Edicions Bellaterra.

Snyder, Joel (2014). *The visual made verbal*. Arlington: American Council of the Blind.

Tor-Carroggio, Irene, and Helena Casas-Tost (forthcoming). "Who is Currently Audio Describing in China? A Study on the Profile of Chinese Audio Describers". To appear in *MonTI. Monografías de Traducción e Interpretación*.

Wu, Zongyi, and Zhenzhen Xie. 2015. 中国视障口述影像服务的发展现状与大陆推广 [Current Development of the Audio Description Service in China and its Popularization in the Mainland]. *Journal of News Research* 6(10): 20–22.

Zhang, Chao. 2017. “‘Nothing about Us without Us’: the Emerging Disability Movement and Advocacy in China”. *Disability & Society* 32:7, 1096–1101, doi:10.1080/09687599.2017.1321229

Zhou, Xun. 2002. “The Discourse of Disability in Modern China”. *Patterns of Prejudice* 36(1): 104–112.

Chapter 4. Article 3: Chinese AD guidelines

When East Meets West: A Comparison of Audio Description Guidelines in China and Europe¹⁷

Resumen

Aunque China es el país con más personas con diversidad funcional del mundo, tanto la oferta como el estudio académico de los servicios de accesibilidad sensorial que permiten acceder a los medios en igualdad de condiciones se encuentran todavía en un estado muy incipiente. El presente estudio parte de dos objetivos: presentar una muestra inédita de tres guías de buenas prácticas de audiodescripción chinas cuya función es, sobre todo, formar a audiodescriptores; y compararlas grosso modo con documentos equivalentes europeos para ver si existen diferencias o similitudes relevantes y lecciones de las que ambas partes se puedan nutrir. Nuestros resultados muestran que existen muchos puntos en común y que la principal diferencia es cómo se aborda la cuestión de la objetividad. Más concretamente, Europa se aleja de la dicotomía entre la objetividad y la subjetividad, mientras que China todavía defiende que la primera es característica indispensable de cualquier buena audiodescripción.

Palabras clave: traducción audiovisual, accesibilidad a los medios, audiodescripción, guías europeas, guías chinas

Abstract

Despite China being the country with most disabled people in the world, the provision of media accessibility services and their academic study are still in their infancy. The present study set out with two objectives: to present a sample of three Chinese audio description guidelines that are mainly employed for training purposes, and to compare the Chinese situation in terms of guidelines with that of Europe, to see if there are any relevant differences to point out and any mutual lessons to be learnt. Our results show there are a number of general similarities between the different Chinese and European guidelines and that the main difference is the different approach to objective vs. subjective description. While Europe is moving away from the dichotomy between objectivity and

¹⁷ Tor-Carroggio, I., & Vercauteren, G. (2020). When East meets West: A comparison of audio description guidelines in China and Europe. *HIKMA. Translation Studies Journal*.

subjectivity, China still advocates objectivity as being a necessary characteristic of quality audio description.

Keywords: audiovisual translation, media accessibility, audio description, European audio description guidelines, Chinese audio description guidelines

4.1. Introduction

China is the country with the most people with disabilities in the world (Wu & Xie, 2015) but, paradoxically, media accessibility is still in its infancy.¹⁸ It is a fact that most Chinese TV programmes and films are captioned, but the reason behind this is more of a linguistic rather than an accessibility-related nature: the wide range of Chinese dialects requires captions to guarantee that everybody can follow the content that is broadcast (Li & Looms, 2015; Casas-Tost & Rovira-Esteva, 2018: 33). Zheng (2017) even argues that captions have more of an aesthetic function in many entertainment, cultural and children’s programmes in order to attract the viewers’ attention. Yet China is not oblivious to the needs of people with hearing loss. Indeed, “[p]ersons with hearing impairments are quite well served at national level with open captions/subtitles with the exception of live content such as news. Sign language interpretation is available to a more limited extent, mainly at local and regional level” (Li & Looms, 2015: 270).

Audio description (AD), however, is not available on Chinese TV. This stands in stark contrast with, for example, Europe, where AD is being offered on TV in various countries. In the UK “most major broadcasters (like the BBC, Channel 4 and Sky) provide AD on 20 per cent of their programmes” (RNIB, 2019), and Germany has been offering AD on the TV since 1993 (Reviere, 2016: 239).¹⁹ In China AD has been offered in films through live sessions and recorded CDs since the early 2000s, and more recently through online platforms such as Ximalaya FM and Youku (Tor-Carroggio & Casas-Tost, 2020). Yet, the provision of the service is still far from ideal and many challenges are yet to be

¹⁸ In this article, when we talk about “China” we are mostly referring to the People’s Republic of China.

¹⁹ For a more detailed overview of the availability of AD on television in various European countries, consult the final report of WP1 of the ADLAB project, downloadable from the Deliverables in the ‘Project’ section on www.adlabproject.eu

responded to, such as the lack of financial resources and the irregular provision of the service across the country (Li & Pan, 2013).

Similarly, the academic study of AD in China is still in its infancy and the literature mentioning it is very limited. Only a handful of scholars refer to the situation in Mainland China (Yeung, 2007; Li & Pan, 2013; Leung, 2015; Wu & Xie, 2015; Feng, 2018), and most of them only do so superficially. One of the reasons for this could be that accessibility-related studies still lack a proper descriptive framework that contextualises the service and provides a clear picture of how it is offered in China. Notwithstanding, there have been some recent efforts to take a more comprehensive snapshot of AD in China. For example, Tor-Carroggio and Casas-Tost (2020) have drawn the profile of Chinese audio describers and Tor-Carroggio (2020) has investigated Chinese users' satisfaction and experience with AD. Yet, the existing Chinese AD guidelines have not been subject to study thus far, although they have taken a special role in improving the quality of the service. A comparison with Europe shows that, again, the situation is completely different there. As Orero (2012: 196) notes:

[t]he issue which has attracted the most attention [in media accessibility studies in Europe] is the analysis of existing guidelines, focusing on the research needed to draft new research-based guidelines or standards, with an eye to using them for training at university level.

Another reason that could explain this interest is that standards are a prerequisite for laws, since a law is meaningless if there is no standard to refer to (Matamala & Orero, 2018: 142). Having said that, it would be advisable to make a terminological remark, since standards are different from recommendations or guidelines. In fact, there are Chinese AD guidelines, but no official standard has been agreed on yet.

This paper aims to contribute to the recent wave of interest in Chinese AD by providing a new perspective from which to analyse the current state of the service, namely that of existing guidelines. More specifically, the objective of this study is twofold. First, to present three unpublished Chinese AD guidelines that mainly serve for training purposes. Second, to compare the Chinese situation in terms of guidelines to that of Europe, to see whether there are any relevant differences to point out and any mutual lessons to be learnt. This comparison was deemed necessary and potentially fruitful given the fact that Chinese AD has developed relatively free from European influence, and vice versa. Since both China and Europe share one same final goal in this regard – offering a quality service

to people with sight loss –, we believe this comparison is a natural step to see where each region stands and what lessons they can learn from each other.

This paper is divided into five sections. The first revises the current situation of European AD guidelines; the second describes the methodology followed in the study; the third presents three different Chinese AD guidelines; the fourth discusses the differences and commonalities between the Chinese and European cases; and the last one draws some conclusions, underlines the limitations of the study and provides some ideas regarding what other paths of research could be explored in the future.

4.2. European Guidelines and Standards

Compared to China, Europe has a much longer-standing tradition when it comes to creating guidelines for AD. Dating back to 1997, Benecke and Dosch published *Wenn aus Bildern Worte Werden. Ein Handbuch für Filmbeschreiber*, the German guidelines that formed the blueprint for many later documents that offer guidance to audio describers. In the first decade of this century, various other countries created their own guidelines: in 2001, the UK's communications regulator Ofcom published their *Guidelines on the provision of TV access services*, an updated and reviewed version of which was presented in 2006; in 2005, the Spanish Association for Standardisation and Certification (AENOR) created the UNE 153020 norm *Audiodescripción para personas con discapacidad visual. Requisitos para la audiodescripción y elaboración de audioguías*, and in 2008 France got its *Charte de l'audiodescription* (Morisset & Gonant, 2008).²⁰ Generally speaking, all these guidelines discuss the same categories of information, namely content selection or 'what to describe', AD style or 'how to describe', timing or 'when to describe' and more or less individual specific issues.

In terms of content selection, all guidelines agree that four constituents should be described, namely the actions, the characters performing and the time and place in which they take place. In addition to all this visual narrative information, most of the guidelines point out that some (intradiegetic) sounds that cannot readily be identified by the target audience should also be included in the AD, just like opening and ending credits and other types of text on screen, including logos, written messages such as indications on screen

²⁰ A more comprehensive overview of different sets of guidelines and a comparison of their contents can be found in Greening, Petré, and Rai (2010) and in Perotti and Valero Gisbert (2017).

of when and/or where a certain movie or scene is set, and subtitles (OFCOM, 2001: 5).²¹ One of the few elements where the existing European guidelines do not seem to be in agreement, is as to when characters should be named. The German guidelines suggest that characters should only be named when their name is also mentioned in the film (unless they are historical characters or their name is only given later in the film). The Ofcom guidelines, on the other hand, suggest that for practical reasons the name of a character can be given as early as possible, unless it should remain a secret for the plot (Greening, Petré & Rai, 2010: 6).

When it comes to the style of the AD, they all contain a few specific instructions on how to describe, but none of them are really exhaustive.²² In summary, the main recommendations advise using fluent and simple language (bearing in mind the AD will have to be voiced later), adapting the description to the genre and the pace of the film, using vivid and varied language including adjectives and adverbs that allow the audience to create its own image, and avoiding patronising and subjective descriptions.

Finally, all the guidelines seem to be in agreement regarding the timing of the AD: generally, descriptions cannot interfere with dialogues and relevant sound effects. The German guidelines go a bit further and state that description should only be added when the production is completely silent (Benecke & Dosch, 2004: 20), but since that never really happens, descriptions can go over (parts of) sound effects that are not essential for the story.

These guidelines have proved to be invaluable for the practice of AD, which would not be at the level it is today without them. However, from a translation studies perspective, they all have some common shortcomings: nearly all of them were drawn up by practitioners who learnt the tricks of the trade on the job.²³ As such, they are fairly intuitive and anecdotic, based on personal experiences, and – probably more problematic from an academic point of view – they lack a solid methodological foundation. We are well aware of the fact that none of these guidelines were ever designed as an academic

²¹ Subtitles are a “unique” category of information and the technique of including them in the audio description is usually referred to as audio subtitling or AST. See for example Remael (2012a, b) for more information on audio subtitling.

²² There are two exceptions: first, the comparison made by RNIB (Greening, Petré, and Rai 2010) brings together all the guidance on how to describe from the different guidelines and as such presents a more extensive overview; second, in the Portuguese guidelines (Neves 2011), which were created in 2011, a full chapter is devoted to the language of AD.

²³ Again with the exception of the Portuguese guidelines created by Neves (2011).

publication, so the foregoing is by no means meant to criticise them. Nonetheless, as pointed out by Vercauteren (2016: 74): “if we claim that audio description is a form of translation, it only seems reasonable that we try to apply paradigms used for other types of translation to audio description too”. The paradigm suggested by Vercauteren (2016) is that of functionalism. It would go beyond the scope of the present paper to explore this choice in detail, but two elements that are relevant in the light of guideline creation are that functionalism requires translators a) to give a detailed analysis of their source text to determine what to translate and how, and b) to decide what strategies they will use to create their translation. These two steps in the translation, c.q. AD process formed the basis of the guidelines developed by the ADLAB project (Remael, Reviere & Vercauteren, 2015), which differ from earlier guidelines in two essential respects. First, they fit the entire AD creation process within the theoretical framework of narratology, which offers audio describers a generally applicable approach to determine and decide what and how to describe.²⁴ Indeed, narratology defines the various basic elements or constituents that can be found in any type of narrative (see for example Bal (1997), Herman and Vervaeck (2005) or Herman (2009) for a comprehensive overview of those basic elements). In other words, narratology explains how stories are created and as such offers audio describers a tool to analyse their source text and determine what narrative elements can be included in their description. On the other hand, narratology also explains how audiences process and interpret stories: it offers insights into what elements audiences really need to understand and follow a story and what elements are less crucial. Therefore, it presents audio describers with a tool to decide what elements should be included in the AD and what elements can be left out if there is not enough time to describe everything. This last part in particular is absent from earlier guidelines, which recommend prioritising some information over others, without explaining how to do that.

A second element that sets the ADLAB guidelines apart from earlier documents, is that it steers clear of the discussion between objective and subjective description, by pointing out that “no one ever sees the same film [...] and this is no different for the blind and visually impaired audience since it is just as heterogeneous as the sighted one.” (Remael, Reviere & Vercauteren, 2015: 16). Therefore “AD too is always subjective to some extent

²⁴ The fact that the (audio-)visual products that are described usually tell a story and narratology – as the theory of story-telling – offers a suitable framework to analyse and explain the audio description process and product, has been discussed before; see among others Kruger (2009, 2010), Vercauteren (2016, 2012) or Vercauteren and Remael (2015).

since it is based by [sic] the interpretation of the audio describer.” (Remael, Reviers & Vercauteren, 2015: 16). So rather than prescribing that descriptions have to be objective or subjective, the guidelines advise that the audio describer has to try and find: “a balance between a personal interpretation and personal phrasing (subjectivity) and more text-based interpretation and phrasing (objectivity) that leaves room for further interpretation by the blind and visually impaired users.” (Remael, Reviers & Vercauteren, 2015: 16).

To achieve that aim, the ADLAB guidelines offer various strategies that audio describers can choose from, ranging from the more objective “her eyes open wide” to the more subjective “she is amazed”, or a combination of these: “her eyes open wide in amazement” (Remael, Reviers & Vercauteren, 2015: 17). This principle also brings the AD process closer to that of other forms of translation, that are never ‘or – or’ but always a matter of degree depending on the specific strategies that are chosen.

4.3. Methodology

A descriptive analysis of the Chinese guidelines that were collected served as the methodological backbone of our study. This analysis was complemented by data gathered in interviews with the people who drafted—alone or within a team—the guidelines under study:

- Mrs. Wu Rina (乌日娜): Head of the volunteer team of audio describers at the China Braille Library in Beijing. Interview sent through Wechat on 16 April 2019.
- Mr. Yu Jiang (于江): Head of the volunteer team of audio describers at the Sound of Light Barrier-free Film & TV Culture Development Center. Interview sent through email on 29 April.
- Mrs. Li Shuying (李淑莹): Head of the “Feeling Images through your Heart” programme (心聆感影) in the Sun Yat-sen Library of Guangdong Province. Interview sent through email on 22 April 2019.

The comparison between Chinese and European guidelines and standards was carried out following a list of items that had previously been agreed on and that allowed us to see what information was included or omitted:

- What to describe (i.e. content selection);

- How to describe (i.e. formulation or AD style);
- When to describe (i.e. timing of the AD);
- Voicing

4.4. Chinese AD Guidelines

The lack of unified Chinese guidelines comes as no surprise since AD is still at its outset in the country. In fact, not even the most widespread audio-visual translation modality, i.e. subtitling, can be evaluated by comparing the output to any national standard (Casas-Tost & Rovira-Esteva, 2018). Matamala and Orero (2018: 150) highlight that it is of the utmost importance that the end users' voice is heard and that their opinion is duly taken into account. Unfortunately, China is still far from defining any unified guidelines related to media accessibility partly because the Chinese users' voice is weak or fragmented. In fact, China lacks unified AD guidelines because AD providers have just started to discuss this issue and have not considered it to be a priority so far. Yet, a proposal of empirically tested AD guidelines has recently been tabled. Drafted by Leung (2018) as part of her PhD thesis, it is based on empirical research carried out in Hong Kong. Although Leung's guidelines offer a response to the criticism that some media accessibility guidelines and standards are not supported by experimental research involving users (see, for example, Orero, 2005), they are still largely unknown in the Mainland, where most of the time, non-tested in-house guidelines are used to train volunteer audio describers.

This section will present three unpublished AD guidelines. Although they were drafted in different Chinese cities, they are not exclusively representative of these, since it is possible that more than one set of good practices coexist in one same city.

4.4.1. China Braille Library (Beijing)

In 2011, the China Braille Library created a specific centre for AD that is responsible for AD research, production, training and activity-planning. It offers weekly sessions of films with live AD in Beijing and it also sells—or often donates—CDs with recorded AD to other AD providers in China (Tor-Carroggio & Casas-Tost, 2020). The China Braille Library audio describes approximately 50 movies every year and records the AD of 20-30 movies on CD. Unlike what happens in many Chinese user associations, the people working in this centre are not all volunteers, since some were employed for their background in film studies or broadcasting. Although its activity mainly focuses on audio

describing films, they also claim to have offered AD in art exhibitions, magic shows, variety shows and TV series.

Their guidelines are used in their training sessions, which are organised around ten times every year and are offered to audio describers from all around China for free. These guidelines defend three basic principles, which, and also according to our interviewee, are all rather unspecific: first, AD must be produced in Standard Chinese; second, not all kinds of audio-visual products can be described; and third, audio describers should not take excessively long pauses. As for the materials that are “audio describable”, our interviewee claimed that those programmes or movies that are too quick or that do not have long enough pauses (such as the news) are not suitable to be described.

The China Braille Library’s guidelines also specify some language requirements that need to be met. For example, language needs to be objective, concise, accurate and formal, but specialised terminology and slang need to be avoided. Furthermore, a clear distinction of what is essential and what is secondary must be made so that not all details are described. Moreover, audio describers must construct complete sentences and avoid AD overlapping with the original dialogues. The guidelines also offer an AD script sample.

4.4.2. Sun Yat-sen Library (Guangzhou)

The AD guidelines used in the Sun Yat-sen Library were also drafted for training purposes. Mrs. Li not only organises training sessions at the Sun Yat-sen Library, where 15 described movies are offered live every year, but also in other libraries and user associations in the province of Guangdong. nto

The author of the guidelines started taking notes on her experience as an audio describer in 2014, and in 2017 she combined all her observations in a Power Point presentation, which she now uses to teach and present her guidelines. She is currently drafting the guidelines in a more conventional format so that they can be shared more easily among the audio describers. The guidelines are updated every time she holds a training session, after which more examples are usually added. Mrs. Li was trained by a Hong Kong AD expert who, in turn, had been trained in the USA, and she has read foreign guidelines, mainly *Pictures Painted in Words. ADLAB Audio Description Guidelines*. What Mrs. Li believes must be exported to China from Western guidelines is the concept of objectivity. Yet, and according to our interviewee, Chinese people still seem to be used to audio describers acting more as commentators. As for other domestic guidelines, our

interviewee admits having read those of the Beijing Hong Dan Dan volunteer organisation. Her guidelines were not drafted with the help of people with sight loss but every time she gets feedback from users she incorporates it into her materials.

These guidelines begin by defining the object of study, their target audience (both primary and secondary) and the various delivery modes that are possible. Yet, it must be stated that the only one carried out on their premises is live due to copyright restrictions. The presentation also enumerates the steps to produce AD, the first of which is to select the material to be audio described. It is recommended that this be carried out taking into account the age and the educational level of the audience. It is also suggested that, in addition to new and popular films, suggestions made by users are adopted as well. At this point, the guidelines offer specific advice for AD beginners and advise against selecting martial arts and science fiction films, if possible. This is suggested so that the intrinsic difficulty of these genres does not discourage volunteers from continuing to collaborate with them. Beginners are also advised not to select movies in which dialogues are scarce and those in which the image does not have excessive relevance. Instead, those films that bear similarities with real life and those in which the story is what stands out the most are recommended. Once the movie is chosen, the source text needs to be analysed and only after that can the script be drafted. The guidelines stress the importance of a quality assurance step after the script is drafted, preferably including end users. Yet, Mrs. Li, admits that, unfortunately, this is not always possible.

After that initial introduction, the guidelines move on to analyse the “when”, the “what” and the “wording and style of AD”. Regarding the “when”, it is indicated that AD can be inserted under the following circumstances: when it does not overlap with the dialogues, when characters appear, when the place or time changes or when action happens. With regards to the “what”, what they call the “principles” of objectivity and neutrality are put forward and their importance is stressed. Emphasis is put on the fact that AD is different from a personal interpretation or a guided appreciation. Indeed, describing subjectively and even venturing to suggest what a character is thinking is still one of the most frequent mistakes for beginners, according to these guidelines. The importance of describing the essence of what happens in the image is also stressed as opposed to providing superficial or secondary details. As for the wording and style of AD, it has to be “clear, simple, direct but vivid and diverse”. The guidelines also point out that the time constraints must be respected, and they recommend the use of short sentences that contain action verbs. They

also recommend the description of elements based on a tactile perspective (e.g. “soft as a pillow”, “thin as a sheet of paper”), and the appropriate use of film terms such as “close-up”, “slow motion” and “flashback”. Word repetitions must be avoided, and the description must suit the film’s style. Finally, it is important not to fill every single gap between dialogues with description.

These guidelines conclude by offering several interesting materials. First, an AD script sample is provided. The script is divided into three columns: the first provides information on what happens in the scene, the second specifies, for example, the speed of the description and the last one includes the description itself. Second, the three “don’ts” of AD according to a user called Zhu Junyi are listed: don’t give too much information, don’t be subjective and don’t disturb the viewing experience. Third, some suggestions on what has to be checked in case of live AD in films are put forward. The document ends with the following message: “People are of the most importance and we need to start from rights. Audio description provides equal opportunities and rights to people with sight loss.”

4.4.3. Sound of Light Barrier-free Film & TV Culture Development Center (Shanghai)

Sound of Light Barrier-free Film & TV Culture Development Center (Guanying zhi Sheng 光影之声, hereinafter SoL) is a Shanghainese association set up in 2016. As Tor-Carroggio and Casas-Tost (2020) reported, it produces about 50 audio described movies every year, which are recorded on their premises. These movies can be accessed at specific facilities, which are all located in Shanghai. Their AD is created by volunteers, who are trained with the help of Mr. Yu Jiang’s guidelines. These training sessions usually last for half a day and end with a test. The guidelines were drafted in 2018 and their main objective is to ensure the quality of their AD scripts.²⁵ As a basis for the guidelines only the Netflix style guide was consulted, since it was thought to be quite simple.²⁶ Yet, Netflix’s stress on objectivity had a great influence on this association. Indeed, SoL had been debating at length whether AD should be objective or subjective at that time, and Netflix made the final decision for them. No other domestic or international

²⁵ Their guidelines were revised in April 2019. This was the version analysed in this study.

²⁶ SoL came into contact with Netflix’s guidelines when this platform contacted them to audio describe the series *Chosen* (2017) into Chinese.

document was used as a reference. Nonetheless, they are aware of the existence of other guidelines, such as the ones drawn up by the Beijing Film Institute.

Given that SoL is led by a person with total sight loss herself, user needs and requirements are faithfully reflected in the guidelines. Also, SoL organises AD screening activities every year to which many users are invited so that they can give feedback on the AD SoL delivers.

Yu Jiang's guidelines start off by defining the basic concepts and explain what elements should be included in the AD. The second section presents what they call "the principles of AD". The importance of objectivity is pointed out because users need to be able to draw their own conclusions based on what is described to them. However, given the ever-present time constraints in AD, the guidelines state that the use of qualitative adjectives such as "handsome" or "adorable", can be used. In terms of when to describe, the guidelines specify that the AD needs to follow the image closely and in an orderly manner but, in general, when there is a pause that lasts between one or two seconds, no description should be added. However, in some cases, if a relevant character appears or something important happens, no matter how short the pause is, it should be used to ensure users follow what is going on. As everything else in these guidelines, this point is illustrated with some descriptions that are only a few characters long (e.g. "in the library", "at KTV with some friends") that are indispensable for the comprehension of the examples provided.

The third section focuses on the requirements for the script. Apart from some layout-related issues, the number of Chinese characters per second (including punctuation) is specified: a maximum of five and a minimum of three. Audio describers also need to be careful not to make the AD too long so that it does not overlap with the dialogue and not too short so that it does not leave the audience unattended and confused. In fact, the guidelines state that audio describers should use their time to the full to depict a scene in the most vivid way. Also, the AD must be specific and correct; therefore, audio describers need to do the necessary research to learn about what they are describing. The guidelines suggest that audio describers should look for the appropriate research materials or consult others.

The fourth section puts forward the language-related requirements. To begin with, everyday language should be used so that people of all ages can understand the AD. Thus,

both overly formal and informal language must be avoided. Moreover, language needs to be formulated in such a way that it can be voiced easily. In order to achieve said objective, the guidelines suggest that audio describers should choose the synonym that makes the text as easy to read as possible and also reminds them to be aware of the tone changes that may arise when combining characters. Given that, for instance, the “s” and the “sh” sound is not always easy to differentiate in Chinese, instead of saying 环视四周 (*huan shi si zhou*, to look around), describers should choose 环顾四周 (*huan gu si zhou*, to look around). Similarly, and since “he” (他, *ta*) and “she” (她, *ta*) are pronounced the same exact way, the use of these pronouns should be avoided in the same sentence when referring to different people. Anachronisms need to be avoided as well. Furthermore, adjectives cannot be used to interpret what is actually shown on the screen. Therefore, if a woman whose face is covered in tears is shown, saying “tears roll down her cheeks” would be more appropriate than saying “she is devastated”. In addition, in some cases it will be necessary to specify what Chinese character we are referring to, since some characters share the same pronunciation. For example, in Chen Jialuo’s *Transfer Student from Taipei* (2012) professor Wang writes a character that is pronounced han on the blackboard. Audio describers should not just say “he writes a han character”, since this does not say much, but “he writes a han character, the *han* from the word *Han chao* [literally, ‘Han dynasty’]”. The explicit reference to a known word will facilitate users making an immediate mental association. Finally, cinematic terms are only allowed when they are essential, when one is completely certain that a word is used in a particular way and, obviously, when the time for the AD allows for that extra information.

The fifth and last section clarifies how to deal with more specific issues, such as:

- What information to include during the opening credits. In this case, the producer, the director, the main actors should be mentioned. Were the describer to have more time, the name of the scriptwriter, for example, can also be provided. The movie can also be presented with a short sentence, always following a similar pattern: “讲述了……的故事” (“the movie tells the story of…”). The time for this depends on how long the initial credits last.
- When to name characters: the guidelines recommend waiting until a character’s name is mentioned in the film to reveal their identity. In order to avoid using a character’s name, some strategies are shared. For example, short phrases like “a

middle-aged woman” or a “bald man” are suggested. Yet, in case there are too many characters, their names can be anticipated.

- How to deal with subtitles: they should either be integrated in the AD or included in the formula “subtitles appear: XXX”. If the subtitles contain the title of the movie, it should be voiced as “The title of the movie appears: XXX”. Also, other text on screen, such as historical background information at the beginning of the film or narrative explanations at the end, should be included in the AD.
- How to deal with foreign languages and Chinese dialects: the guidelines recommend that voice over is used to read subtitles in Standard Chinese. Yet, there is no need to translate English words like “yes”, “no” and “okay” into Standard Chinese because they are generally known. The guidelines also recommend not to choose movies in which the presence of Chinese dialects or foreign languages is too strong to avoid that AD becomes voice-over.
- How to overlap with the dialogues when this cannot be avoided: describers need to look for moments when the dialogue does not provide essential information or when sounds or words are repeated, so that they gain a few extra seconds to include description.
- How to handle music and sounds. Since music and sounds are seen as essential components of the movie, describers should make sure that these can be heard so the audience can appreciate them.

The SoL guidelines are clearly the most extensive ones in China, not only because they touch upon the widest range of issues, but also because they are accompanied by clear examples that illustrate how the suggestions can be implemented. Moreover, the fact that ample input from end users was integrated and foreign guidelines were studied and used as background materials contribute to them being a comprehensive training document.

4.5. Discussion

From the presentation in the previous two sections of this paper, it is clear that there are a lot of similarities between the Chinese guidelines and the ones created in various European countries. In terms of content selection or what to describe, all guidelines are in agreement that the main narrative constituents, i.e. the characters, their (re-)actions and the spatio-temporal setting in which these take place, should – if possible – be included

in the description. In terms of timing or when to describe, again there is general agreement: descriptions should not interfere with dialogues and/or sound effects that are hard to interpret without any visual context. Also in terms of AD style or how to describe, there is some resemblance: both in Asia and Europe guidelines advise audio describers to use generally accepted language that is at the same time vivid and varied, yet easy to voice and understand. In addition, it is recommended audio describers use short, simple sentences that preferably contain action verbs, and to adapt their descriptions to the genre they are describing.

However, the comparison makes it very clear that the two continents are at different stages in the guideline development process. This is reflected by the fact that the Chinese guidelines are of a more prescriptive nature than the (more recent) European ones. One possible explanation could be that AD is still relatively new in mainland China and is still a service that is almost exclusively provided and taught by practitioners and user organisations. In Europe, AD is increasingly being offered as an academic course at universities and other higher education institutions, which also started seeing it as an object of research. Since AD has from its onset been considered as a type of audio-visual translation, this meant that paradigms and principles from translation studies have been applied to AD too. The development echoes Jankowska's (2015) remark that "the guidelines or standards existing in many countries [...] are often established with reference to the personal experience of their creators or to similar guidelines used on foreign markets which are often derivative of other sources" (24). It can also explain the shift that can be witnessed, particularly in the latest documents such as the ADLAB guide, from prescriptive guidelines to more descriptive strategies that offer the describer more freedom.

The most important consequence of this development and arguably the biggest difference between the European and Chinese guidelines, is that Europe is moving away from the dichotomy between objectivity and subjectivity. Rather than presenting a black and white "either-or" choice,²⁷ current European guidelines offer audio describers the choice between various strategies on a continuous scale, ranging from highly objective to highly subjective. This approach acknowledges the very complex nature that is inherent to the AD process, and allows audio describers to create and formulate their descriptions

²⁷ In fact, it is not even a choice, since most of the prescriptive guidelines tell audio describers to use objective language.

depending on the genre of the programme or film, the specific moment or scene at hand, the narrative context, the target audience, etc. This does not mean that Europe advocates an “everything goes” approach to AD. In that respect, Leung’s observation (personal communication, 9 August 2018) that AD in mainland China is sometimes more of a spoken commentary than a description, is valid. AD has to follow the narrative that is being developed and should contain (all) the elements that allow the visually impaired audience to reconstruct that narrative just as the sighted audience does. But, as with any form of translation, there are various ways to achieve that goal.

An interesting step in the development process of the Chinese guidelines – and one that has often been neglected in Europe – is that they are heavily based on end user input. Indeed, in the Chinese case user satisfaction and opinion are taken into account and, subsequently, the guidelines are constantly modified. In fact, Tor-Carroggio’s (2020) paper reporting the results of SoL’s latest questionnaire to gather data on users’ views on the current AD provided exemplifies the interest in seeking users’ feedback that will later be reflected in the guidelines.

Finally, just as in Europe, the importance of guidelines for Chinese audio describers is out of the question. As Tor-Carroggio (2018) showed, 80% of the respondents to her study admitted that guidelines are important to them, which suggests that most Chinese describers have a great awareness of the AD process and find it important to follow some recommendations. In any case, there was a point on which all of them agreed: they were all willing to collaborate with other associations and institutions and share their experience.

4.6. Conclusions and implications of the study

The aim of the present paper was to present different AD guidelines that are being used in Mainland China and to compare them both to each other and to similar documents existing elsewhere, particularly in Europe. As became clear from section 4, there are a lot of general similarities between the different Chinese and European guidelines, particularly in terms of content selection and timing, and to a lesser extent also in terms of AD style. The main difference between Chinese guidelines on the one hand and the most recent European documents on the other, is the different approach to objective vs. subjective description, a difference that can be ascribed to the fact that the former

guidelines are created by practitioners with input from end users, whereas the latter are the product of academic research.

Our study is not exempt from limitations, the main one concerning both the small sample of written Chinese guidelines analysed and the limited number of AD providers interviewed. This is considered to be a limitation due to the size of the country and the several groups of volunteers offering AD independently. Nonetheless, and in light of what was outlined above, interesting avenues for further research open up, such as the comparison of existing ADs to see to what extent different guidelines lead to different types of AD, and/or the testing of different types of AD – based on different AD strategies – to gauge audience preferences and mental effort. Indeed, recent research (Fresno, N., Castellà, J., & Soler Vilageliu, O., 2014; Fresno, N., Castellà, J., & Soler Vilageliu, O., 2016) seems to indicate that different AD strategies lead to a different mental workload induced in the audience, and experimental studies testing different types of AD could shed some initial light on these hypotheses. The results from these tests can then be used to train audio describers and hence further improve AD both on a qualitative and quantitative level. It would also be interesting to compare the way AD is conducted on the Mainland as opposed to Hong Kong, where they already have experiment-based guidelines.

Despite the limitations mentioned, this study has paved the way for more studies that will contribute to the creation of a new academic field -media accessibility- in Mainland China. More specifically, it will set the course for future experimental research that can serve as a solid base for a future Chinese AD standard that can be referred to when the corresponding law is finally passed. Also, we hope this paper will lead to further professionalization and both qualitative and quantitative improvement of AD services in the country. Finally, through the comparison of the Chinese and European traditions, it is likely that the paper will generate discussions that may lead to further adaptations and optimizations of AD as is currently offered in Europe.

Acknowledgements

This paper has been funded by the EasyTV project (GA761999) and the RAD project (PGC2018-096566-B-I00). Irene Tor-Carroggio is member of the TransMedia Catalonia

research group (2017SGR113) and is currently enrolled in the PhD Programme in Translation and Intercultural Studies at the Autonomous University of Barcelona.

References

Bal, M. (1997). *Narratology: an introduction to the theory of narrative*. Toronto: University of Toronto Press.

Benecke, B. and Dosch, E. (1997). *Wenn aus Bildern Worte Werden. Ein Handbuch für Filmbeschreiber*. München: Bayerischer Rundfunk.

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

Casas-Tost, H. and Rovira-Esteva, S. (2018). A quest for effective and inclusive design of Chinese characters in subtitling. *International Journal of Asian Language Processing*, 28(1), 31-47.

Feng, J. (冯纪元). (2018). 口述影像的版权规制探究 (Koushu Yingxiang de Banquan Guizhi Tanjiu) [Study on the Copyright of AD]. *Journal of Henan Science and Technology*, 3, 36-38.

Fresno, N., Castellà, J. and Soler Vilageliu, O. (2014). Less is more. Effects of the amount of information and its presentation in the recall and reception of audio described characters. *International Journal of Sciences: Basic and Applied Research*, 14(2), 169-196.

Fresno, N., Castellà, J. and Soler Vilageliu, O. (2016). ‘What should I say?’ Tentative criteria to prioritize information in the audio description of film characters In A. Matamala & P. Orero (Eds.), *Researching audio description. New approaches* (pp. 143-167). Palgrave Macmillan.

Greening, J., Petré, L. and Rai, S. (2010). A Comparative Study of Audio Description Guidelines Prevalent in Different Countries. Available from: http://audiodescription.co.uk/uploads/general/RNIB._AD_standards.pdf [last accessed: 27 July 2019]

Herman, D. (2009). *Basic elements of narrative*. Oxford: Wiley-Blackwell.

- Herman, L. and Vervaeck, B. (2005). *Handbook of narrative analysis*. Lincoln - London: University of Nebraska Press.
- Jankowska, A. (2015). *Translating audio description scripts: translation as a new strategy of creating audio description*. Frankfurt: Peter Lang.
- Kruger, J-L. (2009). The translation of narrative fiction: impostulating the narrative origo. *Perspectives: Studies in Translation Theory and Practice*, 17(1), 15-32.
- Kruger, J-L. (2010). Audio narration: re-narrativising film. *Perspectives: Studies in Translation Theory and Practice*, 18(3), 231-249.
- Leung, D. (2015). Audio Description in Hong Kong. In R. Baños and J. Díaz-Cintas [Eds.]. *Audiovisual Translation in a Global Context* (pp. 266-281). London: Palgrave Macmillan.
- Leung, D. (2018). Audio Description of Audiovisual Programmes for the Visually Impaired in Hong Kong (PhD thesis). University College London, UK.
- Li, D. (李东晓) and Looms, P. (2015). Television Accessibility in China. In P. Biswas, C. Duarte, P. Langdon and L. Almeida [Eds.]. *A Multimodal end-2-end approach to access* (pp. 261-273). London: Springer-Verlag.
- Li, D. and Pan, X. (2013). 绘声绘色：中国无障碍电影的发展现状及展望 (Huishenghuise: Zhongguo wuzhang'ai dianying de fazhan xianzhuang ji zhanwang) [Vividness: State of the Art of Audio Described Films in China and Their Future Prospects]. *Zhejiang Academic Journal* 4, 188-198.
- Matamala, A. and Orero, P. (2018). Standardising Accessibility: Transferring Knowledge to Society. *Journal of Audiovisual Translation* 1, 139-154.
- Morisset, L. and Gonant, F. (2008). La Charte de l'audiodescription. Available from: www.csa.fr [last accessed: 8 July 2019]
- Neves, J. (2011). Guia de audiodescrição. Imagens que se ouvem. Leiria: IPL.
- OFCOM. (2001). Guidelines on the provision of television access services. Available from: http://audiodescription.co.uk/uploads/general/RNIB_AD_standards.pdf [last accessed: 8 July 2019]

Orero, P. (2005). Audio Description: Professional Recognition, Practice and Standards in Spain. *Translation Watch Quarterly* 1, 7-18.

Orero, P. (2012). Audio Description Behaviour: Universals, Regularities and Guidelines. *International Journal of Humanities and Social Science*, 17 (2), 195-202.

Perotti, O. and Valero Gisbert, M. (2017). Standards for audio description. Contrastive analysis. In A. Bécart, V. Merola and R. López-Campos Bodineau [Eds.]. *Audiovisual translation. Current issues and challenges* (pp. 67-75). Sevilla: Editorial Bienza.

Remael, A. (2012a). Audio description with audio subtitling for Dutch multilingual films: Manipulating textual cohesion on different levels. *Meta: Journal des traducteurs*, 57 (2), 385-407.

Remael, A. (2012b). Media Accessibility. In Y. Gambier and L. van Doorslaer [Eds.]. *Handbook of Translation Studies Volume 3* (pp. 95-101). John Benjamins Publishing Company.

Remael, A., Reviere, N. and Vercauteren, G. (2015). *Pictures painted in words. ADLAB audio description guidelines*. Trieste: Edizione Università di Trieste.

Reviere, N. (2016). Audio Description Services in Europe: An Update. *Journal of Specialised Translation* 26, 232-247.

RNIB. (2019). Television. Available from: <http://rnib.org.uk/television>. [Last accessed: 10 June 2019].

Tor-Carroggio, I. (2018). State-of-the-art of AD standards in China. Poster presented at Languages and the Media, Berlin, 3-5 October 2018.

Tor-Carroggio, I. (2020). The customer is always right: study on Chinese persons with sight loss' opinion on their experience with audio description. *Disability & Society*. doi: 10.1080/09687599.2020.1713727

Tor-Carroggio, I. and Casas-Tost, H. (2020). Who Is Currently Audio Describing in China? A Study on the Profile of Chinese Audio Describers. *MonTI*, 12, 78–107.

Vercauteren, G. (2012). A narratological approach to content selection in audio description. Towards a strategy for the description of narratological time. *MonTI*, 4, 207-231.

Vercauteren, G. (2016). A translational and narratological approach to audio describing narrative characters. *TTR*, XXVII(2), 71-90.

Vercauteren, G. and Remael, A. (2015). Spatio-temporal settings. In A. Maszerowska, A. Matamala and P. Orero [Eds]. *Audio description. New perspectives illustrated* (pp. 61-80). Amsterdam, Philadelphia: John Benjamins Publishing Company.

Wu, Z. (吴宗艺) and Xie, Z. (谢桢桢). (2015). 中国视障口述影像服务的发展现状与大陆推广 (Zhongguo shizhang koushu yingxiang fuwu de fazhan Xianzhuang yu dalu tuiguang) [Current development of the audio description service in China and its popularization in the Mainland]. *Journal of News Research*, 6.

Yeung, J. (2007). Audio Description in the Chinese World. In J. Díaz-Cintas, P. Orero and A. Remael [Eds.]. *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 231-244). Amsterdam: Rodopi.

Zheng, X. (郑玄). (2017). Estándares de subtitulación de la televisión china: Un estudio de caso (MA thesis). Universitat Autònoma de Barcelona, Spain.

**Chapter 5. Article 4: Reception study on the acceptance of
TTS AD**

T(ime) T(o) S(tart) synthesising AD in China? Results of a reception study²⁸

ABSTRACT

Text-to-speech audio description has proven to be an accepted method to increase the output of this access service in several languages, such as English, Japanese, Polish, Spanish and Catalan – at least as an interim solution until more audio description is available with human narrators. In the light of China’s need to make audio description more widely available, we propose text-to-speech audio description as a means to provide this access service both faster and more economically. This article describes and analyses the results of a reception study carried out in China to test the acceptance of text-to-speech audio description in comparison with standard human-voiced audio description. The study sample consisted of forty participants and was carried out using clips from a Chinese historical movie. With the exception of comprehension, the results show that natural voices score statistically higher than synthetic voices, which suggests that those with sight loss prefer human-voiced audio description. Yet, it cannot be stated that text-to-speech audio description is not accepted: more than half of the study participants not only see this as an interim alternative, but also as a welcome permanent solution if it means more audio described movies.

KEYWORDS

Audio-visual translation, media accessibility, audio description, text-to-speech, speech synthesis, reception study.

5.1. Introduction

China is the country with the highest number of people with disabilities (Wu & Xie 2015: 21). It not only numbers among the countries with the oldest populations, along with

²⁸ Tor-Carroggio, I. (2020). T(ime) T(o) S(tart) synthesising AD in China? Results of a reception study. *JoSTrans: The Journal of Specialised Translation*.

Japan, the United Kingdom and the USA (Wang *et al.* 2014: 76), but also its elderly population is expected to rise to 30% by 2050 (Wang *et al.* 2014: 76). Countries facing ageing societies need to tackle a plethora of pressing challenges, including how to accommodate people with disabilities because disability is closely related to age (Peng *et al.* 2010). Although audio description (AD) is still in its infancy in China, especially as an object of academic study, it is of paramount importance to make it more widely available and broaden its scope, as it is mainly restricted to films (Tor-Carroggio & Casas-Tost, *forthcoming*); this would not only benefit people with sight loss, but also sighted audiences (Ramos 2015). Various low-cost proposals have been put forward to extend this type of intersemiotic translation and tested with reception studies, such as AD translation (Jankowska 2015; Jankowska, Milç & Fryer 2017) and text-to-speech (TTS) AD (Szarkowska 2011). The latter is, for example, widely used in theatres that offer AD in Spain (Hermosa, *forthcoming*). According to Szarkowska (2011: 145), TTS AD:

[...] offers unequalled cost-effectiveness in terms of AD production in comparison with conventional methods of producing audio description as it does not require the recording of the AD script (for pre-recorded AD) nor does it incur any human labour costs for the reading out of the AD script (for live AD).

Nonetheless, and unlike languages such as Catalan (Fernández-Torné & Matamala 2015) and Japanese (Kobayashi *et al.* 2010), TTS AD has never been formally tested in Chinese, which means that it cannot yet be considered a potential solution to expedite the production of AD and make it more cost-effective. In fact, TTS AD in Chinese may produce different results from previous research in terms of reception, because TTS is constantly being improved and has to meet different challenges compared to other languages. Wang *et al.* (2006) list some of these challenges, the most salient examples being the large number of commonly used Chinese characters (around 20,000 excluding ancient documents), phonological changes occurring in tonal languages, words not being formally segmented and the polyphony of around 600 characters. Consequently, and in order to study how Chinese AD users respond to the way TTS tackles these challenges, a study was deemed necessary to bridge this research gap. Since TTS AD had already been evaluated in other languages and contexts, a replication study was considered to be the most suitable methodological approach.

Replication is crucial for scientific progress and, as Olalla-Soler (2019) argues, the scarcity of available literature on this issue in Translation and Interpreting Studies begs

the need for more replication studies. Contrasted hypotheses should be tested many times to counteract the high probability of producing false research results (Ioannidis 2005), and in the case of this study, a constructive replication has been performed, i.e., the reliability of the effect detected in previous studies has been assessed by modifying a limited number of aspects of the experimental design (Olalla-Soler 2019).

Hence, the objective of this article is twofold. First, it attempts to evaluate TTS AD in Chinese by comparing it with standard human-voiced AD regarding key features. Second, it also attempts to analyse and demonstrate whether TTS AD would be accepted in China as either an interim solution until there is more AD available with human narrators, a permanent solution, or both. This article is divided into four sections. The first reviews previous studies on TTS, particularly TTS AD, in order to situate this research initiative within the current academic context. The second describes the methodology employed for this quasi experimental research. The third presents and discusses the results obtained and the final section sets out the study conclusions, its limitations and future perspectives to explore.

5.2. Previous research on TTS

TTS has proven to be an effective support instrument for people with sight loss (Cryer & Home 2008). Its applications are quite diverse and range from global positioning systems (GPS) to educational, communication and entertainment tools (Cryer & Home 2008: 5). Some studies suggest that how acceptable these voices are depends on how much experience users have with them (Szarkowska & Jankowska 2012), and that not only is it possible to get used to them (Hjelmquist, Jansson & Torrell 1990 in Cryer & Home 2008), but also in some cases users have stated they prefer these types of voices. The reasons given are their lack of expressiveness, because this feature allows them to concentrate more on the content, it grants them more autonomy and guarantees more confidentiality (Llisterri *et al.* 1993). Although in most cases users prefer natural voices (Szarkowska 2011; Fernández-Torné & Matamala 2015), they are also aware that products with artificial voices expedite access to information, since they do not need to wait for someone to voice the content for them. This was clearly shown in the study by Thompson *et al.* (1999 in Cryer & Home 2008), whose objective was to discover how those with sight loss preferred to receive tax information. Stevens *et al.* (2005) noted that the key factor behind accepting synthetic voices lies in how natural they sound. In short, these previous studies

suggest that users' subjective evaluations of artificial voices are based on how much experience they have with them, the context in which they are used and the voice features.

Cryer and Home (2009) conducted a study on the acceptance of TTS in audio books. This study is important as it is the only one which collected very valuable qualitative information that can be extrapolated to other immersive fields such as movies. In their case, the opinions reported by users were quite varied: some saw the potential of the idea, while others considered it to be an aberration and stressed its shortcomings. However, most participants came to the conclusion that not all types of books were suited to being read with artificial voices: while likely applications were instructional, educational or non-fiction books, a human narrator was preferred for fiction and leisure reading. Cryer and Home (2009: 26) also reported that the acceptance of TTS audio books is very likely to be experience-dependent, since TTS is better accepted among people who have been previously exposed to it. Yet, they also admitted that this is not always the case. Finally, despite the diverse opinions, there were a few participants in their study who stated they would accept this solution if it meant more access to information.

5.2.1. Previous research on TTS in AVT

TTS has been applied to various modalities of AVT, namely voice-over, audio subtitling (AST) and AD. Regarding the former, Matamala and Ortiz-Boix (2018) compared TTS with human voices in a wildlife documentary with a sample of 16 participants and two clips that lasted around two minutes. Using questionnaires, they elicited participants' opinions in terms of self-reported interest, engagement and enjoyment. Participants were also asked to rate the quality, naturalness and comprehensibility of the voices they heard, as well as to respond to questions on comprehension and preferences. Although human voices were rated higher in general, there were no differences in their self-reported engagement. However, a clear positive opinion about using artificial voices as a possible substitute for human voices for this type of translation was not observed, although more than half of their respondents stated excerpts voiced with TTS would be acceptable in a TV broadcast.

TTS AST has also been tested and is actually a reality in many countries such as Spain (Rovira-Esteva & Tor-Carroggio 2018), the Netherlands, Denmark and Sweden (Iturregui-Gallardo 2019). While some authors focused on providing improved technical

solutions to deliver TTS AST (Derbring *et al.* 2009), Thrane (2013) evaluated the difficulties encountered by 16 TTS AST consumers when watching the news, documentaries and fiction content. She concluded that some of the most problematic areas were synchronisation, pronunciation, the presence of multiple voices and speed. She also discovered that TTS ASTs were considered more acceptable for non-fiction content such as the news. A similar conclusion was reached by Kobayashi *et al.* (2010), who found that TTS AD was generally accepted, especially in short informative videos. Kobayashi *et al.* (2010) assessed the acceptance and effectiveness of TTS AD in various genres through surveys and in-depth interview sessions in both Japan and the US. In the survey distributed in Japan, users ranked the voices heard (human, prototype TTS and standard TTS), whilst during the interviews they were asked to evaluate their experience using enjoyability and intelligibility as evaluation criteria. Although the study carried out in the US shared the same objectives, it was slightly different. The survey was designed to test comprehension and comfort, as well as effectiveness and preferences. The test ended with in-depth interviews to explore the characteristics of TTS AD.

Kobayashi *et al.*'s (2010) study has not been the only one putting TTS AD to the test. Previous studies also explored TTS AD acceptance in Polish through questionnaires using a wide range of materials: an educational animated series (Walczak 2010), a Polish feature film (Szarkowska 2011), a foreign film dubbed into Polish (Drożdż-Kubik 2011), a foreign movie with voice-over (Szarkowska & Jankowska 2012) and a documentary with AST (Mączyńska & Szarkowska 2011). Despite the variety of genres evaluated, the conclusions were similar in all cases: TTS AD is accepted in most cases as an interim solution, and sometimes even as a permanent solution. Nevertheless, TTS AD seems to be more acceptable with non-fiction content. It is worth noting that Szarkowska and Jankowska (2012: 86) pointed out quality (especially intelligibility and naturalness) as playing a crucial role in user comprehension and attitudes towards TTS. Also, and compared to viewers with low vision, Szarkowska and Jankowska (2012) detected that blind participants are more supportive of TTS. This finding could be explained by blind viewers being more dependent on AD and thus wanting more audio described films regardless of the voice.

Besides Polish, TTS AD has also been tested in other European languages, such as Catalan (Fernández-Torné & Matamala 2015). Fernández-Torné and Matamala (2015) tested TTS AD acceptance in a dubbed feature film. Their reception study was conducted

with 67 users, who assessed two synthetic voices applied to AD, as well as two natural voices. Participants were administered a questionnaire to rate voices taking into account various end user reception-related items, and to answer questions about their personal preferences. The conclusion reached is in line with those of the Polish project: most participants accepted Catalan TTS AD as an alternative solution to the human-voiced AD. However, natural voices outperformed the artificial ones tested and were still the preferred solution.

These studies focused mainly on the acceptance of TTS AD, but other aspects have also been investigated, such as the effect of TTS AD both on emotion and presence (Fryer & Freeman 2014; Walczak & Fryer 2017) using the emotion elicitation scale and the ITC-Sense of presence inventory, as well as preference questions. Results indicate that higher levels of presence are obtained for AD delivered by a human voice, especially for drama. Likewise, AD delivered by a human voice also enhances emotion.

Finally, although no academic study of TTS AD in Chinese can be reported, the Chinese company Shanghai Gaozhi Keji Ltd. (上海高智科技公司) had already produced a few films with TTS AD back in 2014. According to this company, the production of an entire film usually took them around three days and they used a free online TTS to voice the scripts (the voice can be listened at https://www.xfyun.cn/services/online_tts). The main problem detected at the time was related to the sentence rhythm, especially in names of people and places that are not well-known. This happened before the current group of radio presenters that volunteer to voice ADs in Shanghai was created (Tor-Carroggio & Casas-Tost, *forthcoming*). Yet, this TTS AD was never tested (formally or informally) with users and was rapidly substituted for human-voiced versions.

5.3. Methodology

The stimuli used in our study came from a fiction product kindly provided by the Shanghainese AD production centre Sound of Light (Guangying zhi Sheng, 光影之声). This non-profit organisation describes 50 movies every year and each AD script is supervised by an AD user. We worked from Tor-Carroggio's study (2020) on Chinese AD user movies preferences, which concluded that informants (N = 52), who were mostly elderly people from Shanghai, preferred historical movies when asked which type they would like this experiment to be carried out with. This genre ranked first and second,

respectively, in the two questionnaires presented. This finding sheds some light on the study participants' preferences, since those recruited for this experiment were expected to have a similar demographic profile. This situation was due to the author having a limited number of user sources, which obviously limits the representativeness of our sample. Consequently, the Chinese historical movie *Our time will come* (Mingyue Ji Shi You, 明月几时有), directed by Ann Hui and released in 2017, was chosen. This movie revolves around the resistance movement during Japan's occupation of Hong Kong in the 1940s.

The film and its AD script were analysed in depth in order to select two clips that were as comparable as possible in terms of content (food is mentioned in both), length (almost 5 minutes), intervening characters (the film's two leading actresses appear in both), background music (almost no music or striking sounds are heard), and AD density (around 600 characters each). This information is summarised in Table 5.1. Both clips start describing what Fang Mu (one of the main characters' mother) is doing and coincide in mentioning items such as a mirror and rice. Both clips include people's names, as this is one of the elements that can be problematic for TTS to resolve. Also, the AD briefly overlaps with the dialogue in both cases. Yet, the spots were wisely chosen by the Chinese audio describers because no relevant information is missed, as the Sound of Light guidelines recommend (Tor-Carroggio & Vercauteren, *forthcoming*). Finally, in these two clips, some of the AD units were rather long. We thought having longer units would allow users to listen to the voice for a longer time and, therefore, get more used to it. This was deemed important as TTS acceptance appears to be dependent on how much users are exposed to it (Szarkowska & Jankowska 2012).

The movie was already described in Chinese with a male human voice, consequently, the gender of the artificial voice was chosen to match the original clip (male as well). The male voice was synthesised using the Chinese software Ke Da Yuyin Ku 4.0 (科大语音酷 4.0) as it is available for free. Since in one of the clips only female characters intervened, having a male voice was considered to be a good choice to be able to clearly differentiate the AD voice. Regarding the audio mix, the researcher was assisted by the company Shanghai Gaozhi Keji Ltd, which has experience in AD mixing. At this point, it must be pointed out that, although TTS AD can be read directly by speech synthesis

software, recording it was seen as the safest option taking into account the number of individual homes the researcher had to visit and the various problems that could arise.

Only one minor change had to be applied to the original AD script so that the TTS would have a better final effect. The retroflex sound *er* (JL), a characteristic of Chinese northern dialects which could be not pronounced in all the cases in this sample without affecting comprehension, was deleted because the TTS would not join its sound with the character just before it, which is how it should be read.

Table 5.1. Clips selected and their characteristics

Clip	Human voice	Artificial voice	Characteristics
Clip A	Male	Male	4'56'' min 556 Chinese characters
Clip B	Male	Male	4'37'' min 592 Chinese characters

A questionnaire to rate the voice in each clip was designed. So, each participant answered two questionnaires, which were administered orally by the researcher. These questionnaires included comprehension questions, as recommended by Chmiel & Mazur (2012), and a list of nine parameters that had to be evaluated after listening to each clip:

- Voice naturalness
- Voice pleasantness
- Speech pauses
- Ease of listening
- Comprehension
- Pronunciation
- Intonation
- Acceptance
- Overall impression

These items were selected mainly from the International Telecommunications Union (ITU) Recommendation P.85 (1994), which defines a testing method for evaluating the

subjective quality of synthetic speech. Other authors were consulted to add other suitable items to be evaluated, such as Hinterleitner *et al.* (2011) and Viswanathan and Viswanathan (2005), following the research designed by Fernández-Torné and Matamala (2015).

Although the objective (i.e. technical) acoustic evaluation of synthetic voices has proven to be useful “to measure where voices differ from a human utterance” (Cryer & Home 2010: 6), a more subjective approach involving user testing was taken for one main reason: users are those who would ultimately use TTS AD, although they “differ in both ability and opinion” (Cryer & Home 2010: 7). Also, “[s]ubjective user testing is more useful for someone considering the voice for use in a product or service, to find out whether users are happy with the voice” (Cryer & Home 2010: 4). The most common test for measuring opinions in this regard —and the one recommended by the ITU— is the Mean Opinion Score (MOS). This test involves participants listening to synthetic speech and rating the voices on a simple 5-point categorical scale; scores are then averaged across the group. In our case, though, the scale was changed to 0-10, which is allowed by the ITU (1994: 1). Other studies have also used scales with higher granularity because, theoretically, they can result in smaller standard deviations of MOS (Streijl *et al.* 2016). In our case, this was done mainly to avoid statistical ties as much as possible.

The experiment was approved by our university’s Ethics Committee in June 2019 and piloted with three users before formally conducting it. In each experiment the informant was first read the information sheet and the consent form out loud. Their expressed consent to participate in the study was recorded on audio. During the experiment each participant watched two clips: one described with TTS and another one described with a human voice. Table 5.2. presents the listening order of the voices for participants, which followed a Latin square. This order was repeated with all the users recruited. Artificial voices (*) were always presented first to avoid a negative impact on their evaluation, as did Kobayashi *et al.* (2010) and Fernández-Torné and Matamala (2015). The latter based their decision on van Santen (1993), and Viswanathan and Viswanathan (2005: 62).

Table 5.2. Listening order of the clips

User	Clip 1	Clip 2
01	A*	B
02	B*	A

A post-questionnaire was drafted based on Szarkowska and Jankowska (2012) and Fernández-Torné and Matamala (2015). It was aimed at gathering more qualitative information and the participants' demographic data. Given that all our participants suffered from sight loss because they are the main target group AD caters for, there were no questions relating to their disability. This was decided in accordance with the capabilities approach suggested by Mitra (2006), which essentially explains how disability may be triggered by three different factors: the individual's personal characteristics (e.g., impairment, age, race, gender), the individual's resources and the individual's environment (social, economic, political). Consequently, it was believed that disability can be disentangled from being closely related to a physical impairment, so questions regarding disability were deemed unnecessary.

All the questionnaires were translated into Chinese by a professional Chinese translator, whose work was checked by another Chinese translator. The translators were required to use rather simple language because we expected that many users would be elderly people with little or no education at all.

Users were recruited through personal contacts and with the kind help of the Shanghai Guide Dogs' Club and the Shanghai Association of Persons with Disabilities. The only requirement was to have watched at least one audio described movie. All the tests took place in Shanghai, more specifically in the users' homes, in community centres and in East China Normal University. The session in East China Normal University was followed by a focus group with ten users that shared their views with us on TTS AD and on AD in China in general.

Finally, the quantitative results were analysed statistically using SPSS (version 25). Non-parametric tests were run (i.e. Wilcoxon signed-rank test and Spearman's rank correlation coefficient) due to the limited size of our sample. The p-value threshold for the declaration of statistical significance was set at 0.05.

5.4. Results and discussion

This section presents the results yielded by our study. It begins with a presentation of the demographic characteristics of our sample and then moves on to present both the quantitative and qualitative data gathered.

5.4.1. Description of the sample

Forty AD users took part in the test, 52.5% of whom were men, while 45% were women (one person did not answer this question). As expected, most of the participants were quite elderly: average 58.64 and median 61 (Min = 30, Max = 78). Regarding their educational background, the majority of the participants had completed secondary education or had attended a vocational school (72.5%), close to a fifth had attended university (17.5%), one person had only completed primary education and another had not received any formal education at all (two participants declined to answer this question). As for their experience with TTS, 85% of the participants claimed to have used it before (Figure 5.1), mainly in their mobile phones and computers, and more than half of the participants reported using TTS often or very often (Figure 5.2).

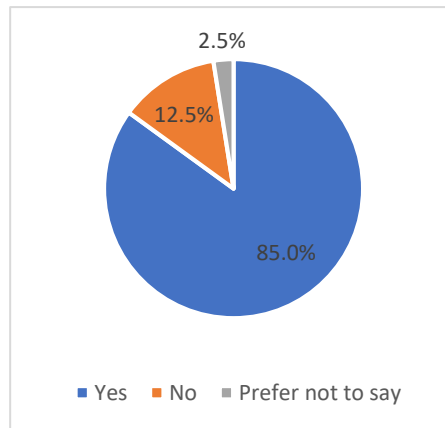


Figure 5.1. Experience with TTS

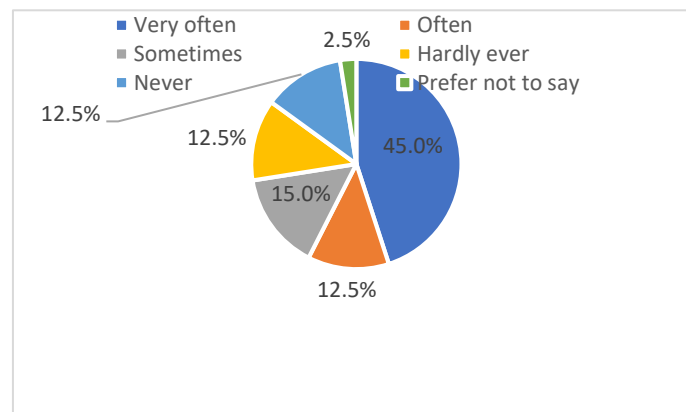


Figure 5.2. TTS frequency of use

All the participants had experience with AD at least once. More than half of the AD users recruited reported using AD in movies often or very often (67.5%) whenever it is

available, which is not always the case. This has to be framed within the Shanghainese context: despite having other opportunities at different venues, only one live audio described movie is available in cinemas every month, screened in specific cinemas but just for one day (Tor-Carroggio & Casas-Tost, *forthcoming*).

5.4.2. Comprehension and quality of the voices

After watching each of the two clips, the users were asked to answer three comprehension questions. The average number of correct answers was 1.42 for the clip audio described with TTS and 1.75 for the clip audio described with a human voice. Yet, no statistical differences were detected ($Z = -1.62$, $p = 0.10$, with 14 ties). These rather low scores might be explained by different factors, such as some of the users being quite aged, which could mean that their cognitive faculties are less sharp. Also, the number of correct answers being higher in the case of the clip voiced with a human voice can be regarded as an effect of the order in which they were presented. Since the TTS version was always presented first, participants may have understood the clip better the second time or been prepared for the comprehension questions. It was decided not to tell them in advance what kind of questions they would need to answer because we intended to simulate a real situation in which users would not be paying excessive attention to every single detail for fear of a test. Furthermore, almost half of the tests were carried out in the users' homes, which were not always quiet environments. Choosing their homes as the testing venue was the only option possible in many cases, since it was quite inconvenient for many of them to leave their homes. However, this situation was seen as an opportunity to see users' response to a more ecological environment, which in the end is where AD TTS would most likely be used were it to be implemented. After carrying out a Mann-Whitney U test, no statistical differences were spotted in terms of correct answers to the comprehension questions between those who watched the clips at home and the rest who did not ($Z = -0.28$ and $p = 0.77$ for the TTS clips, and $Z = -0.23$ and $p = 0.81$ for the clips described with a human voice).

Nonetheless, and in order to clarify which factors had had an influence on the informants' comprehension, we decided to recruit a younger sample to watch the same clips in similar conditions. Fifty sighted third-year undergraduate students agreed to participate. They were BA Translation and Interpreting students from the Hangzhou Foreign Languages Institute, and were blindfolded during the experiment, which took place in one of their

classrooms. In their case, the average number of correct answers was 1.80 for the clip audio described with TTS and 2.10 for the clip audio described with a human voice. Although it might seem that the students understood the contents of the clip better when described with a human voice, no statistical differences were detected ($Z = -1.81$, $p = 0.07$, with 17 ties). These slightly better results might be linked to the students' age (they were all in their early twenties), to their education background and/or to the testing venue environment (which was quieter). Yet, no statistical differences were detected between the students and the AD users either for the first clip ($Z = -1.92$, $p = 0.054$) or for the second one ($Z = -1.62$, $p = 0.10$). This shows that, in this second round, the parameter 'comprehension' did not obtain statistically better results with younger and better-educated respondents in a quieter environment. Yet, it is also true that it was the first time for our participants to listen to an AD, so maybe they found it hard to cope with the initial excitement and so much aural input. In any case, this gives some food for thought regarding the cognitive load AD users can cope with, regardless of their age and education background. In fact, when the questionnaire was first piloted with a user who works as an AD reviewer, this person also got some comprehension answers wrong and underlined the need to review what aspects were important to include in an AD script.

It should be pointed out, however, that although sighted students participated in the study and answered the exact same questions as the users with sight loss, this article mainly focuses on the data obtained from the visually impaired users, since, as Walczak (2010: 39) notes, they are the major beneficiaries of AD. Mendoza and Matamala (2019) reached a similar conclusion in Spain: unlike subtitles, people consuming AD tend to be visually impaired. In fact, these researchers found that not even the professionals of this service usually watch audio described products. Given that AD is still a novelty in China, and also that the students who participated had no previous experience with AD, their participation in the study was purely anecdotal and also used for them to experience the possibilities AD experimental research can offer. Therefore, the following data refers exclusively to the participants with sight loss.

As for the quality of the voices, the average score for each parameter tested was as follows (Table 5.3.):

Table 5.3. Descriptive results

Parameter	MOS TTS	MOS Human voice
Voice naturalness	8.35	9.60
Voice pleasantness	8.20	9.18
Intonation	7.65	9.25
Pronunciation	9.10	9.60
Speech pauses	8.77	9.38
Ease of listening	9.20	9.63
Comprehension	9.80	9.85
Acceptance	9.13	9.68
Overall impression	8.57	9.33

Although the average number of correct answers for comprehension questions was rather low, most users thought they had understood everything. In fact, the parameter that was best assessed in both cases was ‘comprehension’. Yet, what is especially relevant is the fact that all items are assessed above 6, which is the minimum score to pass any test in China and, thus, our study’s threshold. All the parameters related to the human voice were above nine and scored better than those related to the TTS, however, some parameters from the latter also scored above nine (pronunciation, ease of listening, comprehension and acceptance). The parameter that leaves most room for improvement is intonation, but that comes as no surprise since other studies (such as Walczak 2010) had already warned about unnatural intonation being one TTS’s main drawbacks. Similarly, Fryer and Freeman (2014) also concluded prosody is a critical component of AD content in terms of presence and emotion elicitation, and Kobayashi *et al.* (2010: 167) had already reported users complaining about TTS being “less comfortable due to their ‘flat’ intonation.” The importance of this parameter cannot be underestimated since, for example, Ramos (2015: 88) claimed that intonation clearly influences the emotional impact of texts.

Table 5.4. shows that statistical differences were detected in almost all the parameters under study. Thus, it can be stated that the quality of the human voice is regarded as better than that of the artificial one, except for the parameter ‘comprehension’, in which no statistical differences were spotted.

Table 5.4. Wilcoxon test results

Parameter	Wilcoxon test results
Voice naturalness	$Z = -3.29$ $p < 0.01$ ties = 10
Voice pleasantness	$Z = -2.73$ $p < 0.01$ ties = 11
Intonation	$Z = -3.51$ $p < 0.01$ ties = 11
Pronunciation	$Z = -2.29$ $p = 0.02$ ties = 21
Speech pauses	$Z = -2.40$ $p = 0.01$ ties = 21
Ease of listening	$Z = -1.99$ $p = 0.05$ ties = 22
Comprehension	$Z = -1.42$ $p = 0.15$ ties = 32
Acceptance	$Z = -2.04$ $p = 0.04$ ties = 20
Overall impression	$Z = -2.60$ $p < 0.01$ ties = 15

It was deemed necessary to check whether there were any correlations among the parameters tested for the TTS. Statistically significant correlations were found in all cases

except for two pairs ('speech pauses' and 'comprehension', and 'comprehension' and 'acceptance'). The statistically significant correlation coefficients were above 0.5.

5.4.3. Preferences

Apart from quantitative data, our participants also provided some qualitative comments regarding their preferences when consuming AD, as well as the idea of applying TTS to AD. All the comments gathered are reproduced in this section. At this point it should be mentioned that users were generally open to any ideas that could eventually allow them to watch movies along with a sighted audience, for example, using earphones.

When asked about which of the two voices heard they had liked better, half of the participants selected the human voice, although some users were not very satisfied with it because they thought it had an accent and that it seemed to be reciting rather than describing. The comment regarding the voice talent's accent surprised us, since the majority of volunteers currently audio describing in Shanghai are professional radio presenters who are trained to speak perfect Standard Chinese. Future studies should not take this for granted and make sure the voice reflects the standard accent. Yet, 27.5% of those who responded had no clear preference and had liked both, while 20.0% preferred the artificial voice (Figure 5.3). In fact, the artificial voice was unexpectedly praised for being clear and some of those who liked it better claimed that it was very smooth and did not have an accent. Szarkowska and Jankowska (2012: 85) also report many users liking standard accent voices.

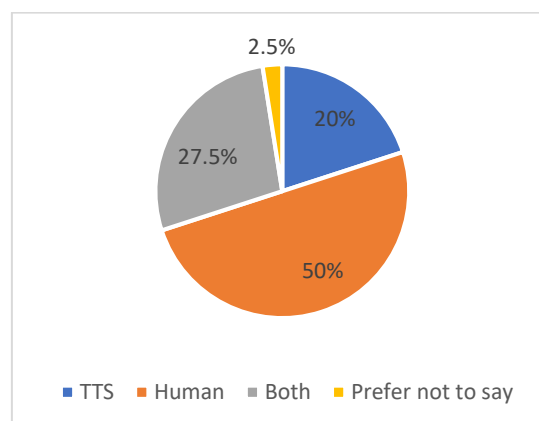


Figure 5.3. Voice preferred

Nonetheless, when asked what type of voice they would like to listen to in audio described films, over 50% chose human voices (57.5%), whereas only one user selected TTS (2.5%).

Despite this clear difference in proportion, 15% of those who responded said they did not care and almost a quarter of the participants (22.5%) claimed that it depended on the movie. Those who stated that it depended on the type of movie considered documentaries to be the most suitable genre for TTS AD. This coincides with Kobayashi *et al.* (2010: 170), who concluded that TTS AD is more “suitable for informational videos where understanding is the critical factor”, and Fernández-Torné and Matamala’s findings (2015). Similarly, Walczak and Fryer (2017: 77) found that the users recruited preferred to watch dramas with a human voice and were more sceptical about doing so with TTS. For the documentary, “out of 36 participants, 31 (86%) were eager to watch the documentary also with TTS AD” (Walczak & Fryer 2017: 77). Similarly, Walczak and Fryer (2017) came to the conclusion that presence rates, as well as levels of interest and confusion were comparable for documentaries audio described by a human voice and TTS. Yet, it should be noted that documentaries are not the products that benefit most from AD (Walczak 2010: 43). Our participants also considered historical and martial arts films to be appropriate for TTS AD. There was also one user who underlined AD was an art and the reason why TTS AD is not suitable for all kinds of movies. Nevertheless, users felt being able to understand the AD was more important than the voice used.

Although there was a clear preference for human voices (both quantitatively and qualitatively), all participants except for one stated that they would accept TTS AD as an interim solution until there is more AD available with human narrators (Figure 5.4) and almost 63% of the users would even accept it as a permanent solution (Figure 5.5).

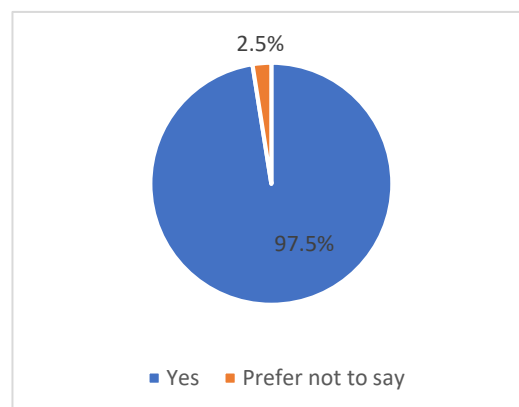


Figure 5.4. Interim acceptance

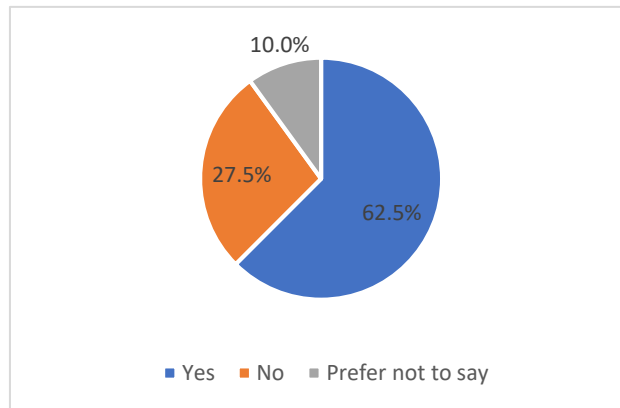


Figure 5.5. Permanent acceptance

As Table 5.5. illustrates, these numbers are very similar to those obtained in previous studies.

Table 5.5. Comparison with similar studies

Study	Language and audio-visual stimuli	Acceptance of TTS as an interim solution	Acceptance of TTS as a permanent solution
Szarkowska (2011)	Polish/ Polish feature film	95%	58%
Szakowska and Jankowska (2012)	Polish/ Foreign film with Polish voice-over	95%	70%
Fernández-Torné and Matamala (2015)	Catalan/ Miscellaneous film	94% believe TTS AD can be an alternative solution to human-voiced AD	
Present study	Chinese/ Historical film	97.5%	62.5%

The acceptance of TTS AD in the Chinese context cannot be divorced from the scarcity of audio described movies, which means that users do not have enough movies to choose

from. Many of our participants complained about not being able to choose what audio described movies to watch and also about the scantiness of newer and foreign films. Szarkowska and Jankowska (2012: 84) had already pointed out that people with sight loss want to watch foreign films just like the sighted audience. Therefore, they were open to solutions that could increase the availability of audio described movies. Nonetheless, not everybody who accepted TTS AD did so because the alternative was fewer ADs or none at all. Some of our participants had difficulty in differentiating the artificial voice from the human one, and some did not even realise that the first clip had been described with an artificial voice. Also, some participants stated that having real voices for movie AD did not necessarily guarantee AD excellence. In fact, some users confessed checking who the AD voicer is in the monthly live AD sessions in cinemas to see whether they like the volunteer in charge of voicing AD script or not. Moreover, some participants mentioned that it is sometimes difficult to know when the actors talk and when the describer describes, since the voices might be very similar. These issues could be solved by using TTS AD, but also by making a better choice of human voices or a better mix of the original soundtrack with the AD.

Two users also mentioned that TTS is constantly improving, so they were already sure TTS AD could substitute traditional AD in the near future. In fact, some of them were aware of new TTS systems cloning human voices, which will make it necessary to replicate this study to validate our current findings. These two same users also raised an interesting point: they were willing to pay for AD if that meant having more choices and having it online so that they could watch audio described movies at home. They also highlighted the fact of not liking having to feel grateful to the volunteers all the time and having to restrain themselves from pointing out aspects they do not like about the AD currently being delivered.

Notwithstanding, the participants mainly complained that TTS AD was too monotonous and did not convey any emotions. But again, this was seen as a less relevant problem in comparison with not having access to more audio described materials.

Finally, it was deemed necessary to see whether the TTS acceptance item was correlated to either the participants' frequency of use of TTS or to how often they consume AD. No statistically significant correlations were found in such cases ($r = 0.06, p = 0.72 / r = 0.08, p = 0.61$). Yet, the users who participated in the focus groups mentioned them not having difficulties in accepting AD TTS for them being very used to artificial voices.

5.5. Conclusion

This article has presented the results of a reception study carried out in China to test the acceptance of TTS AD in comparison with conventional AD. The results show that natural voices have statistically higher scores than synthetic voices, which proves that people with sight loss prefer AD voiced by a human. Yet, this does not mean that TTS AD is not accepted, since it is viewed as both an interim alternative and even as a permanent solution by more than half of the participants if that means more access to audio described movies. We have also confirmed that prosody is still a pending subject for TTS, compared to other parameters. Therefore, if Chinese TTS companies were to suggest their voices for AD, they should improve them in this regard. Furthermore, the fact that our results confirm previous studies can be attributed to AD's situation in China being equivalent to that of, for example, Poland's ten years ago. Our participants gave the impression they would accept anything that could increase AD availability.

Although it is not the first time that TTS AD has been tested with users, our study has incorporated a series of features that make it innovative. To begin with, TTS AD had never been tested in Chinese. In fact, no AD-related tests had ever been conducted in Mainland China, so our study hopes to encourage other researchers, especially in the Chinese academic world, to engage in media accessibility research and, more specifically, in user testing. Moreover, the clips used in the study are slightly longer than those used by Fernández-Torné and Matamala (2015). This was something Fernández-Torné and Matamala suggested as a way of taking TTS AD research a step further. Fernández-Torné and Matamala (2015) also suggested testing it with different genres, which this study has also contributed to. In addition, this study followed a user-centric design and the genre was selected by users in a pre-study. Furthermore, comprehension questions were added before asking the users' opinion as regards the voices used. This was done to assess user comprehension, since it is usually taken for granted but is crucial and has priority over technological innovations. We found that the use of TTS AD did not have a statistically significant impact on comprehension.

Obviously, this research has its limitations and we would like to note seven observations. First, only one synthetic voice is tested, which is also only compared to one natural voice. Furthermore, these two voices were not selected by AD users, so it is uncertain whether voices chosen by them would have yielded different results. Future studies should put to

the test other voices that have previously been validated by users. Second, given the size and the socioeconomical disparities in China, our results cannot be extrapolated to the whole of China. Third, the size of our sample, which is not random, leaves room for improvement in future studies. Ramos (2015) acknowledges the difficulty of recruiting volunteers with sight loss. According to Ramos (2015: 87), “[t]his justifies the fact that reception studies in audio-visual translation usually work with smaller samples.” Fourth, although the clips were slightly longer than those used by Fernández-Torné and Matamala (2015), they are still short and, therefore, studies with complete movies audio described with TTS should be conducted in the future. Fifth, Fernández-Torné and Matamala (2015) also suggest testing for engagement, which would have been useful in this case as well. Sixth, we are aware that by following the design by Fernández-Torné and Matamala (2015), the order of the clips was not sufficiently randomised (the synthesised speech was always presented first). The order effects could be further explored but right now our choice could also be seen as a limitation. Seventh, since we did not ask about the type of disability, it would have made more sense to include more potential AD users, such as people with mental impairments. Nevertheless, it is hoped that this study has provided an empirical basis from which to initiate further debate and contributions to this area.

Despite incorporating these innovative aspects, and given AD’s dependence on volunteers, it would be interesting to investigate other ways to produce AD in a less time-consuming way in Chinese. An example of this could be AD translation, which has already proven to be successful in some language pairs, such as English and Polish (Jankowska 2015), while Szarkowska (2011) also suggested exploring AD templates. These two ideas could be particularly useful in Shanghai, since voicers outnumber AD scriptwriters. Also, it would be helpful to study how TTS can be used to provide more didactic tools for Chinese children with sight loss and how TTS affects emotion and presence rates.

Acknowledgements

This paper has been funded by the EasyTV project (GA761999), the Confucius Institute and the Department of Translation and Interpreting and East Asian Studies of the Autonomous University of Barcelona (UAB). The author is member of the TransMedia Catalonia research group (2017SGR113) and is currently enrolled in the PhD Programme

in Translation and Intercultural Studies at UAB. The author would like to thank the users who participated in the study and the reviewers for their enriching and helpful suggestions.

References

- **Chmiel, Agnieszka and Iwona Mazur** (2012). “AD reception research: some methodological considerations.” Elisa Perego (ed.) (2012). *Emerging topics in translation: Audio description*. Trieste: EUT, 57-74.
- **Cryer, Heather and Sarah Home** (2008). “Exploring the use of synthetic speech by blind and partially sighted people. Literature review #2.” Birmingham: RNIB Centre for Accessible Information (CAI).
- **Cryer, Heather and Sarah Home** (2009). “User attitudes towards synthetic speech for Talking Books.” Birmingham: RNIB Centre for Accessible Information (CAI).
- **Cryer, Heather and Sarah Home** (2010). *Review of methods for evaluating synthetic speech*. Technical report #8. Birmingham: RNIB Centre for Accessible Information (CAI).
- **Derbring, Sandra, Peter Ljunglöf and Maria Olsson** (2009). “TS: Light-weight automatic reading of subtitles.” Kristiina Jokinen and Eckhard Bick (eds) (2009). *NODALIDA 2009 Conference Proceedings*. Odense: Northern European Association for Language Technology 272–274.
- **Drożdż-Kubik, Jan** (2011). *Harry Potter i Kamień Filozoficzny słowem malowany – czyli badanie odbioru filmu z audiodeskrypcją z syntezą mowy*. Masters dissertation. Jagiellonian University.
- **Fernández-Torné, Anna and Anna Matamala** (2015). “Text-to-speech vs. human voiced audio descriptions: a reception study in films dubbed into Catalan.” *The Journal of Specialised Translation* 24, 61-88.
- **Fryer, Louise and Jonathan Freeman** (2014). “Can you feel what I’m saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment.” Anna Felnhofer and Oswald D. Kothgassner (eds) (2014).

Proceedings of the International Society for Presence Research. Wien: Facultas, 99-107.

- **Hermosa, Irene** (forthcoming). “Delivery Approaches in Audio Description for the Scenic Arts.” To appear in *Parallèles*.
- **Hinterleitner, Florian et al.** (2011). “An evaluation protocol for the subjective assessment of text-to-speech in audiobook reading tasks.” *Proceedings of the Blizzard Challenge Workshop*. International Speech Communication Association (ISCA). Florence.
- **Hjelmquist, Erland, Bengt Jansson and Gunilla Torell** (1990). “Computer-oriented technology for blind readers.” *Journal of Visual Impairment and Blindness* 17, 210-215.
- **Ioannidis, John PA** (2005). “Why most published research findings are false.” *PLoS Medicine* 2(8), 696–701.
- **ITU-T Recommendation P.85** (1994). *Telephone transmission quality subjective opinion tests. A method for subjective performance assessment of the quality of speech voice output devices*. Geneva: ITU. <http://www.itu.int/rec/T-REC-P.85-199406-I/en> (consulted 7.11.2019).
- **Iturregui-Gallardo, Gonzalo** (2019). *Audio Subtitling: Strategies for Voicing Subtitles and Their Effect on Film Enjoyment and Emotion*. PhD thesis. Autonomous University of Barcelona.
- **Jankowska, Anna** (2015). *Translating Audio Description Scripts: Translation as a New Strategy of Creating Audio Description*. Berna: Peter Lang.
- **Jankowska, Anna, Michał Milc and Louise Fryer** (2017). “Translating audio description scripts... into English.” *SKASE Journal of Translation and Interpretation* 10(2), 2-16.
- **Kobayashi, Masatomo et al.** (2010). “Are Synthesized Video Descriptions Acceptable?” *ASSETS '10: Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*. New York: ACM, 163-170.

- **Llisterri, Joaquim et al.** (1993). “Testing users’ acceptance of Ciber232, a text to speech system used by blind persons.” Björn Granström, Sheri Hunnicut and Karl-Erik Spens (eds) (1993). *Speech and Language Technology for Disabled Persons. Proceedings of an ESCA Workshop*. Stockholm: KTH-ESCA, 203-206.
- **Matamala, Anna and Carla Ortiz-Boix** (2018). “Text-to-speech voice-over? A study on user preferences in the voicing of wildlife documentaries.” *SKASE Journal of Translation and Interpretation* 1, 24-38.
- **Mączyńska, Magdalena and Agnieszka Szarkowska** (2011). “Text-to-speech audio description with audio subtitling to a non-fiction film “La Soufriere” by Werner Herzog.” Paper presented at the *Media for All Conference. Audiovisual Translation: Taking stock* (Imperial College London, 28 June 2011).
- **Mendoza, Nuria and Anna Matamala** (2019). “Panorama de la enseñanza de la audiodescripción en España: resultados de un cuestionario.” *MonTI. Monografías de Traducción e Interpretación* 11, 155-185.
- **Mitra, Sophie** (2006). “The Capability Approach and Disability.” *Journal of Disability Policy Studies* 16(4), 236-247.
- **Olalla-Soler, Christian** (2019). “Practices and attitudes toward replication in empirical translation and interpreting studies.” *Target* 32(1), 3-36.
- **Peng, Xiaoxia et al.** (2010). “Ageing, the urban-rural gap and disability trends: 19 years of experience in China - 1987 to 2006.” *PLoS ONE* 5(8), e12129.
- **Ramos, Marina** (2015). “The emotional experience of films: does audio description make a difference?” *The Translator* 21(1), 68-94.
- **Rovira-Esteva, Sara and Irene Tor-Carroggio** (2018). “Serveis d'accessibilitat sensorial a les televisions que emeten en català: situació actual i propostes de futur.” *Quaderns del CAC* 44(21), 71-80.
- **Stevens, Catherine et al.** (2005). “On-line experimental methods to evaluate text-to-speech (TTS) synthesis: effects of voice gender and signal quality on intelligibility, naturalness and preference.” *Computer Speech and Language* 19, 129-146.

- **Streijl, Robert C., Stefan Winkler and David S. Hands** (2016). “Mean Opinion Score (MOS) revisited: Methods and applications, limitations and alternatives.” *Multimedia Systems* 22(2): 213–227.
- **Szarkowska, Agnieszka** (2011). “Text-to-speech audio description: towards wider availability of AD.” *The Journal of Specialised Translation* 15, 142-163.
- **Szarkowska, Agnieszka and Anna Jankowska** (2012). “Text-to-speech audio description of voice-over films. A case study of audio described *Volver* in Polish.” Elisa Perego (ed.) (2012). *Emerging Topics in Translation: Audio description*. Trieste: EUT, 81-98.
- **Thompson, Leanne, Christopher Reeves and Kate Masters** (1999). “In the balance: making financial information accessible.” *British Journal of Visual Impairment* 17(2), 65-70.
- **Thrane, Lisbeth** (2013). *Text-to-speech on Digital TV An Exploratory Study of Spoken Subtitles on DRISyn*. Masters dissertation. University of Copenhagen.
- **Tor-Carroggio, Irene and Helena Casas-Tost** (forthcoming). “Who Is Currently Audio Describing in China? A Study of Chinese Audio Descriptor Profiles.” To appear in *MonTI* in 2020.
- **Tor-Carroggio, Irene** (2020). “The customer is always right: study on Chinese persons with sight loss’ opinion on their experience with audio description.” *Disability & Society*. Doi: 10.1080/09687599.2020.1713727
- **Tor-Carroggio, Irene and Gert Vercauteren** (forthcoming). “When East Meets West: A Comparison of Audio Description Guidelines in China and Europe” To appear in *HIKMA* in 2020.
- **van Santen, Jan P.H.** (1993). “Perceptual experiments for diagnostic testing of text-to-speech systems.” *Computer Speech & Language* 7(1), 49–100.
- **Viswanathan, Mahesh and Madhubalan Viswanathan** (2005). “Measuring speech quality for text-to-speech systems development and assessment of a modified mean opinion score (MOS) scale.” *Computer Speech and Language* 19, 55-83.

- **Walczak, Agnieszka** (2010). *Audio description for children. A case study of text-to-speech audio description of educational animation series Once Upon a Time... Life*. Masters dissertation. University of Warsaw.
- **Walczak, Agnieszka and Louise Fryer** (2018). “Vocal delivery of audio description by genre: measuring users’ presence.” *Perspectives* 26(1), 69-83.
- **Wang, Ren-Hua et al.** (2006). “Mandarin text-to-speech synthesis.” Chin-Hui Lee *et al.* (eds) (2006). *Advances in Chinese Spoken Languages Processing*. Singapore: World Scientific, 99-124.
- **Wang, Hua et al.** (2014). “Zhongguo Renkou Laolinghua Shehui Fazhan yu Yingdui Celüe.” *Zhongguo Shehui Yixue Zazhi* 31(2), 75-77.
- **Wu, Zongyi and Zhenzhen Xie** (2015). “Zhongguo shizhang koushu yingxiang fuwu de fazhan xianzhuang yu dalu tuiguang.” *Xinwen Yanjiu Shikan* 6(10), 20-22.

Biography

Irene Tor-Carroggio is a Ph.D. student in Translation and Intercultural Studies at the Universitat Autònoma de Barcelona (UAB) and is also a member of the research group TransMedia Catalonia (2017SGR113). She holds a B.A. in Translation and Interpretation from the UAB (2013) and also an M.A. in International Business from Shanghai University of Finance and Economics (2017). She is part of the EU-funded project EasyTV, <http://easytvproject.eu>.



Email: irene.tor@uab.cat

Chapter 6. Discussion and conclusions

This concluding chapter presents part of a discussion that could not be included in the corresponding publication due to space limitations. It also addresses the most significant findings of this study as a whole, highlights its limitations, provides possible guidelines with reference to future research and enumerates the activities carried out to share the acquired knowledge and disseminate our results.

6.1. Discussion

Due to the space limitations encountered in Article 1, the specific objective 2a could not be fulfilled in it. This section will specifically focus on the comparison between the data gathered from Chinese audio describers (and already presented in Chapter 2) and that obtained by the ADLAB PRO project, in which European audio describers were profiled.

In the case of the ADLAB PRO project, 65 responses from audio describers were collected. This response rate was slightly higher than ours ($N = 53$). The ADLAB PRO respondents were mostly women (66.2%), like in the Chinese case (75.5%). The ADLAB PRO project recruited participants from a variety of European countries, including Spain, Poland, Slovenia, Italy and the UK.²⁹ Yet, their report mentions that “project partners’ respective countries are best represented” (ADLAB PRO, 2017, p. 9). Something similar has happened to us, with audio describers working in Shanghai featuring more prominently in our study due to us having more contacts in that specific city.

Regarding what AD-related tasks participants carried out (see Figure 6.1. , all the European respondents wrote or revised AD scripts, while not even half of our respondents had experience in doing so. In our case, the majority of respondents were AD voicers (almost 70%) and none had ever translated an AD script, neither by themselves nor by using machine translation and post-editing. This is unfortunate, since, as Jankowska (2015) states, translating AD is a “tempting alternative strategy for creating audio

²⁹ The overall number of responses analysed in the ADLAB PRO project was 183. Responses were collected from audio describers ($N = 65$), AD end-users ($N = 100$) and service providers ($N = 18$). Since the demographic questions were common for all profiles, the results are mostly presented altogether in their report (see ADLAB PRO, 2017). Among the respondents there were 26 persons from non-European countries such as Mexico, Brazil and Canada. Yet, the ADLAB PRO report (2017) does not specify the profile of these non-European persons. Thus, it is unclear whether they were audio describers, users or service providers.

description scripts in those countries where audio description is still scarce”. This is due to AD translation reducing or eliminating the need for several steps necessary to create AD from scratch, such as cuing and consulting with other audio describers or end users. AD translation is already a reality in Europe and is used by, for instance, VOD platforms that produce their own content. Yet, in the light of our results, it can be seen that AD translation is still not the most widespread method to create AD in Europe. In addition, despite AD being in its infancy in China, there are already individuals carrying out AD quality control, although the number of people doing so in Europe is around four times higher.

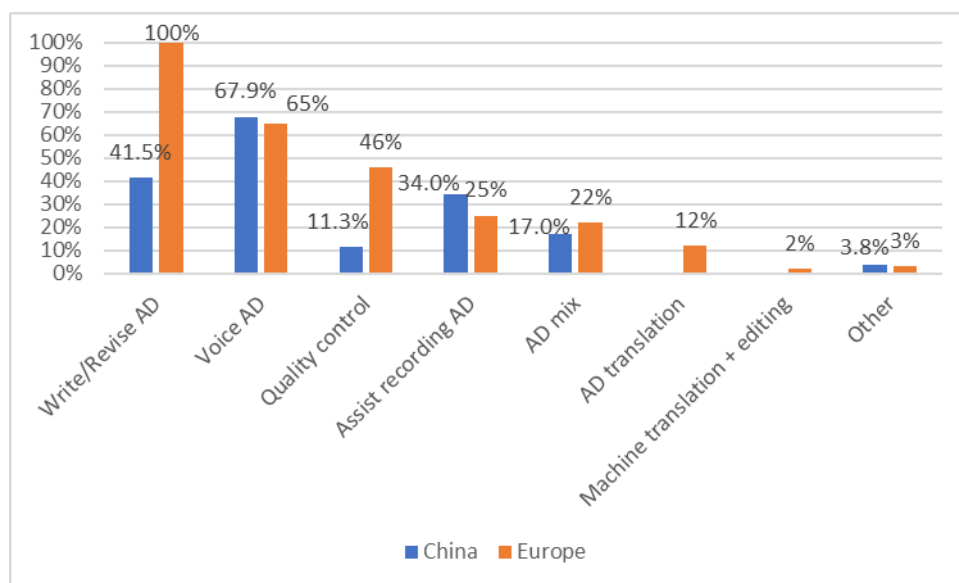


Figure 6.1. Tasks performed by ADLAB PRO and our respondents

Although not many European audio describers claimed to translate AD, around 30% of them had working experience related to Translation Studies (see Figure 6.2.). In fact, the ADLAB PRO results show that “audio describers coming from a different working environment than AD generally have an AVT background since they are former audio-visual translators (e.g. subtitlers, dubbing translators, etc.)”. Yet, the European landscape is quite fragmented. Although the differences with the Chinese case are evident (with Chinese audio describers mostly being professional presenters), there is one similarity: in general, audio describers mainly come from professional fields where the use of language, voice, and writing skills is central.

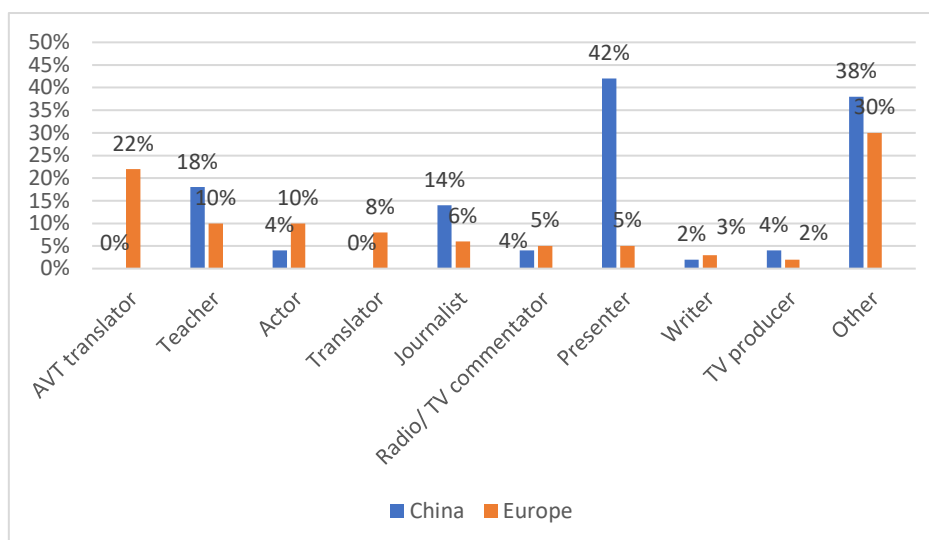


Figure 6.2. ADLAB PRO and our respondents' working experience

These results are in line with the participants' educational background in both cases. In the case of China, the majority of participants had a Film/TV Studies, Journalism/Media Studies and Language/Linguistics background. The European audio describers recruited by ADLAB PRO also had a varied background, but more often than not it was related to the field of humanities and arts, "with a majority of professionals having received education in language and linguistics (40%) or literature (42%). Translation (34%) as well as film and TV studies (28%) are other preferred areas of provenance followed by journalism and media studies (17%)". The results are not surprising in any of the cases, since excellent knowledge of and writing skills in the mother tongue are necessary to work as an audio describer. Yet, it should be pointed out that Translation Studies is already one of the paths audio describers come from in Europe and this stems from AD being considered an intersemiotic type of translation. In China this concept is still very new, and no Translation courses offer AD-training to date in the country.

This conceptual connection is also evidenced when looking at the answers of both the Chinese and the European respondents when asked about what professional figure audio describers resemble. While the majority of European audio describers opted for "audio-visual translator", our Chinese participants chose "presenter/commentator". Apart from the fact that Translation and AD have not been academically connected in China yet, the Chinese participants' answer can be explained by remembering that many of our participants were Shanghai AD voicers working for a radio station.

Regarding the experience accumulated by the participants, most European audio describers have been audio describing for less than ten years (70.8%), whilst in the Chinese case over 80% have five or less years' experience with AD, with 34.0% of them audio describing for less than a year. Therefore, the conclusion reached in the ADLAB PRO project is applicable in our case as well: AD is a rather new discipline both in Europe and China, although in the latter it seems to be even newer. Also, in both cases it stands to reason that the number of experts is rather limited, since only 7.7% of the European audio describers recruited accumulate more than 20 years' experience and just 1.9% of our Chinese participants have been audio describing for over two decades.

With regard to the hours of AD material produced in their career, our sample showed that more than half of our respondents (62.3%) had produced less than 50 hours of AD whilst around half of the European sample had produced over 300 hours. In contrast, only 3.8% of the Chinese informants had produced more than 300 hours of AD.

In Europe, the majority of audio describers are "professional, trained, and paid workers" (ADLAB PRO, 2017, p. 12). This stands in stark contrast with what was found out about our Chinese participants, with most of them willing to produce AD without being professionally involved in this activity. Yet, the ADLAB PRO project also detected some people willing to audio describe without being professionally involved in this field. Despite this difference, in both cases, around three quarters of the participants stated that they had had enough time to satisfactorily complete the AD they have been commissioned with. This was thought to be surprising because it was expected that European audio describers would complain about deadlines, since it is a common complaint in the Translation field. In contrast, this was expected in the case of China because it was assumed that Chinese audio describers would have less time pressure due to them not being economically compensated for their work.

As for the perception audio describers have of AD, both in Europe and China a considerable number of respondents (57.0% in the European case and in 34.0% China) were inclined to believe that AD is both an art (an inherent natural talent) and a craft (something you can learn with training). Nonetheless, the answers of our Chinese respondents were more equally distributed, with opinion divided between those who thought it resembles more an art (28.3%) and those who claimed it is more of a craft (22.6%). Chinese participants considering AD to be a craft is a promising result as it may lead to more individuals pursuing a professional training course in the future.

Audio describers were also asked to rate some statements regarding AD. Despite the distance between European and Chinese participants, a substantial number of similarities were found between the Chinese and the European case. For example, both in Europe and China AD is not considered to be a well-paid job, with 60% of European audio describers agreeing or strongly agreeing with this statement. The percentage of Chinese audio describers disagreeing with this claim reaches almost 91% in the Chinese case. The difference, of course, lies in most Chinese audio describers offering the service for free. As expected, most audio describers, no matter where they come from, subscribe to the idea that AD is socially useful (89% of European audio describers and around 94% in the case of China). The slightly higher proportion of people stating that AD is socially useful could be explained by the Chinese socialist values attached to making society move forward. Yet, there are more European audio describers that find AD satisfying, maybe as it is also a paid job. Over 40% of both Chinese and European audio describers agreed on AD being stressful and around 75% thought AD is a demanding activity. In addition, just over half of both the Chinese and the European audio describers think AD is well-known by users with sight loss, whilst they disagree with AD being popular among the general public. These results are summarised in Table 6.1.

Table 6.1. Statements regarding AD

	Strongly disagree		Disagree		Undecided		Agree		Strongly agree	
	EU	CH	EU	CH	EU	CH	EU	CH	EU	CH
Prestigious	6%	5.6%	22%	13.2%	35%	32.1%	28%	39.6%	9%	9.4%
Well-known by the general public	49%	11.3%	39%	58.5%	6%	16.98%	3%	13.21%	3%	0.0%
Well-known by users with sight loss	2%	0.0%	15%	11.3%	22%	32.1%	45%	47.2%	17%	9.4%
Stressful	3%	7.6%	28%	35.8%	26%	7.6%	38%	32.1%	5%	17.0%
Demanding	2%	0%	2%	11.3%	14%	13.2%	49%	50.9%	34%	24.5%
Well-paid	26%	58.49%	34%	32.1%	26%	5.6%	14%	1.9%	0%	1.9%
Satisfying	3%	1.9%	0%	9.4%	3%	16.9%	46%	49.1%	48%	22.6%
Creative	0%	3.8%	5%	1.9%	2%	1.9%	35%	62.3%	58%	30.2%

Socially useful	1%	5.7%	5%	0%	5%	0%	20%	15.1%	69%	79.2%
------------------------	----	------	----	----	----	----	-----	-------	-----	-------

As far as training is concerned, the survey data shows that while nearly all the ADLAB PRO respondents had received some form of training (especially in film and TV AD), almost 40% of the Chinese audio describers lacked any kind of specific training. It is worth noting that the training received in China is not comparable to the European case because the latter already has years of accumulated experience and can even be found at university level, whilst the training offered in China is rather short-termed and informal. What is more, it would have been interesting to interview Chinese audio describers to gain more insight into types of AD training since, despite AD being relegated mostly to films, some audio describers claimed to have received training to audio describe teaching materials, theatre and other live events, for example.

The majority of European respondents received specific AD training (75%), while the opposite was observed in the Chinese case (64.2%). In-house training (conducted at a company/institution, etc.) was the most popular option in both cases (46% in Europe and around 50% in China). Yet, in Europe the training possibilities are more varied and, unlike China, even include courses at university level. Moreover, half of the European audio describers received a certificate after training, but they are hardly ever asked to show it to obtain a job. In the Chinese case, more than three quarters of the respondents said that they did not receive a certificate and hardly any of them had ever been asked to show any training proof to audio describe. In both cases this indicates that official training is not required yet and it would be interesting to see what the actual requirements are when choosing/employing an audio describer. We believe that in China they cannot be too demanding for two reasons: first, no extensive training is offered and, second, AD is, to this date, a voluntary activity in which many people engage for free, so requirements need to be slightly lowered. Still, we are aware of audio describers needing to meet some basic requirements, such as a good command of their mother tongue.

According to our data, although the number of respondents working mainly alone or always alone is considerable (37.7%), the number of respondents doing so mainly in a team or always in a team is also noteworthy (30.4%). Conversely, in the European case, audio describers always or mainly work alone (see Figure 6.3.).

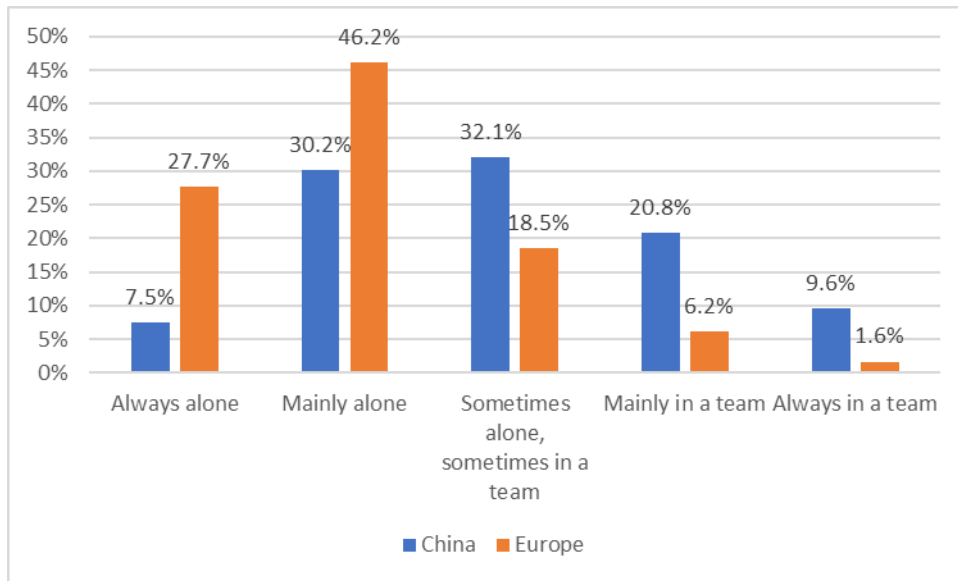


Figure 6.3. How audio describers work

In the ADLAB PRO project it was concluded that AD is a solitary activity, probably due to the advent of new technologies making it possible to carry out AD-related activities from home. We cannot conclude the same in the Chinese case, but it is not surprising given the fact that one of the main AD-related complaints in China is that the service needs to improve technology-wise.

Another difference detected was regarding the cooperation between audio describers and people with sight loss during the AD scripting process. While in Europe audio describers do not seem to attach importance to (or find it difficult to) collaborate with AD users, almost half of our Chinese respondents do so always or sometimes (see Figure 6.4.). This can be explained by the fact that some AD promoters are AD users themselves, as is the case of Sound of Light and the workshop run by Professor Jiang Hongyuan in Shanghai. Since in Europe AD is already a legally protected service, broadcasters, for example, need to make an effort to comply with what the law mandates. In contrast, in China AD is not a legal obligation for any public or private company yet, so the people providing it are usually individuals that are somehow benefitting from it directly (users, users' relatives...) and that want to promote it for personal or ethical reasons. It is hoped that this collaboration is not reduced once AD becomes a legal requirement, since as we already noticed, Chinese audio describers acknowledge how useful it is to work with them.

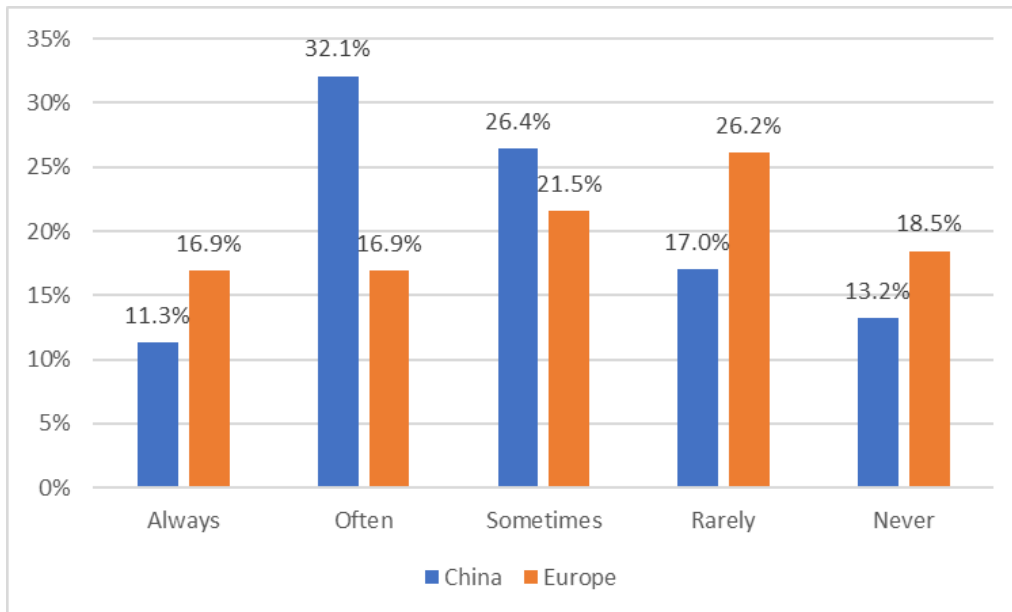


Figure 6.4. Audio describers collaboration with users

According to the ADLAB PRO results, nearly half of the audio describers (42%) only sometimes ask the opinion of other audio describers to overcome specific AD problems (with 23% asking it often and 9% always). Only 26% never or rarely feel the need to talk to their peers. In China, however, almost three quarters of the respondents claimed to often or sometimes ask other audio describers their opinion on how to tackle certain AD issues. In both cases, the comment in the ADLAB PRO report (2017, p. 44) applies: “this indicates a humble approach to the process, the high value given to the opinion of peers, an eagerness to share solutions, and perhaps a way of overcoming the solitary nature of AD work”. As the ADLAB PRO report (2017, p. 44) observes, this can stem from the lack of training materials or guidelines: “the most effective way of learning was and still is by identifying best practices”.

Figure 6.5. shows how audio describers entered the AD field. It is clear that in Europe many audio describers start audio describing after having attended an AVT or specialised course, while Chinese audio describers pursue the profession after an acquaintance introduces them to AD and encourages them to take part in it.

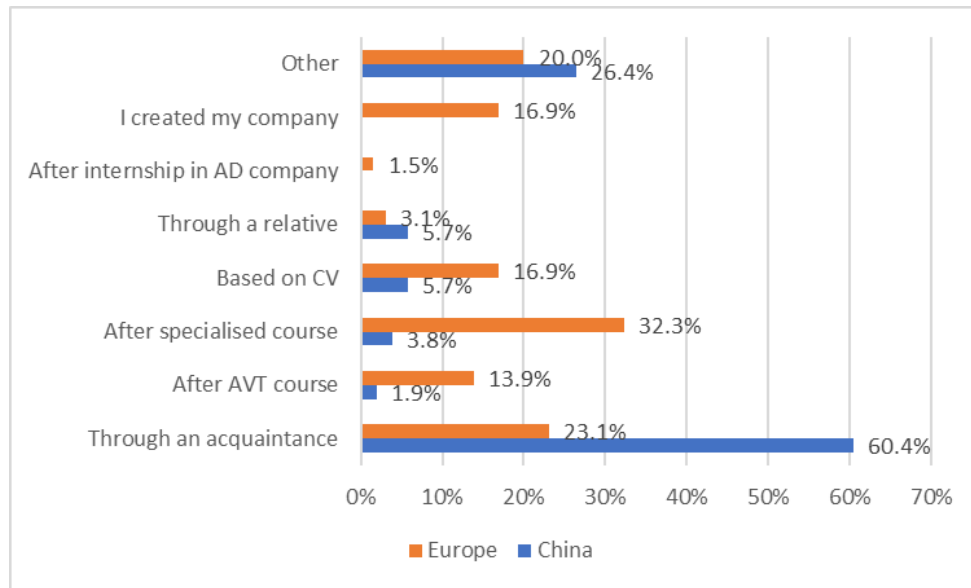


Figure 6.5. How audio describers started working in this field

One of the most striking differences between the ADPALB PRO participants and ours is related to AD guidelines. While almost all European describers were aware of the existence of AD guidelines, this number dropped to 30.2% in the Chinese case. Despite this unawareness, 83.0% of the Chinese respondents thought guidelines were important. This shows the need to draft AD guidelines in China that serve both for training purposes but also as reference for those creating AD. Guidelines are an important reference point for audio describers and therefore attention should be paid in China to present existing guidelines (such as the international ISO standard). Alternatively, local guidelines should be created based on empirical data related to users' needs and preferences.

In both cases, audio describers confirmed their interest in self-improvement (less than 10% of the sample declared that they were currently doing nothing to improve their skills). Experience in the field was considered to be important by the vast majority of Chinese respondents (84.9%). This was also viewed as an effective way of improving one's skills in the European context, but the number of respondents agreeing on the importance of this is far more discrete (37%). No other preferred way to continue to improve AD-related skills and competences was observed. Audio describers are, in general, keen on studying existing material (guidelines, academic articles, etc.) and participating in conferences and workshops, as well as analysing existing ADs in order to learn from other colleagues' work and doing research. Yet, in Europe there seems to be less in-house training offered (25%), whereas in China it seems to be a way to facilitate audio describers' retraining (30.2%). This could be due to China being aware of audio describers lacking a solid basic

training that needs to be compensated for by the association audio describers volunteer for.

6.2. Conclusions

In this PhD dissertation we have explored the past, the present and what could be the near future of AD in China. The first general objective of this dissertation was to describe the history and current state of AD in Mainland China, both as an access service and as an object of study. In order to achieve this aim, four specific objectives were set:

- a) To present the history and the current state of AD in China, both as an accessibility service and as an academic research field, including the legal framework within which this access service is framed.
- b) To profile audio describers in Mainland China.
- c) To investigate the habits, needs and satisfaction of AD users towards this accessibility service.
- d) To present some current Chinese AD guidelines.

As for the first specific objective, one of the main conclusions drawn from this research is that AD is still in its infancy in China, both as a professional practice and as an academic discipline. It has been carried out since the early 2000s, but it is heavily dependent on volunteers and grassroot initiatives and is mostly applied to films. Yet, before finishing this dissertation we have been given more information on what new areas AD is evolving into. For example, Shanghai's Sound of Light informed us that they have recently audio described a theatre play and would like to explore other possible AD applications. As for the legal status of AD, there is national-level legislation mentioning some "narrations" that probably refer to AD, but no specific legal document has been found that clearly advocates for the provision of this access service neither in terms of quantity nor quality. Yet, AD is regularly provided in, for example, first-tier cities such as Beijing, Shanghai and Guangzhou. This access service is also available in other Chinese cities, but its availability is more restricted and sometimes non-existent, especially in the countryside. This is closely linked to China's huge interregional socioeconomic differences.

Regarding the second specific objective, it was found out that the Chinese audio describer's figure is divided into two profiles: AD scriptwriters and voice talents, who are usually embodied by different people. In the light of the results of the study in which a sample of Chinese audio describers were profiled, we can state that the vast majority of respondents were volunteers working on AD part-time or sporadically, did not have broad experience audio describing and many had not received any training. It must be pointed out that, although in our study some audio describers claimed to have been trained in other fields apart from films (such as TV), we must cast doubt on them understanding that specific question correctly, since, for example, no AD is offered on Chinese TV, as far as we know. The Chinese audio describers recruited claimed that they sometimes audio describe alone or in a team although there is a slight tendency towards working alone. Still, almost half of them collaborate with people with sight loss and a high number of our respondents recognised the benefit of working with AD users. Since we had more voicers than AD scriptwriters represented in our sample, it comes as no surprise that many had a Media or Journalism background. It also makes sense that a significant number of respondents was unaware of the existence of AD guidelines since no national AD standard can be found and those consulted in China (always in-house materials) mainly revolve around aspects related to AD writing, not AD voicing. It was also found that many audio describers got involved in AD thanks to an acquaintance and that they have enough time to work on their AD. This is reasonable taking into account that the majority are volunteers working for no apparent reward. Finally, Chinese audio describers stated they are not well-paid and that the AD they produce is not well-known by the general public, which is understandable bearing in mind what we know so far about this access service in China. Furthermore, Chinese audio describers are very aware of the social impact of what they do. We believe this is an especially motivating factor for them, since they live in a society constantly stressing (for example with posters on the streets) the importance of socialist values, including equality.

As for the third specific objective, the analysis of users' habits, needs and satisfaction regarding AD was also one of our thesis' pillars, mainly because the interaction with them was meant to result in information to be used in a reception study that truly involved them and gave a response to their needs. The users that completed our questionnaires confirmed the need of AD and, therefore, validated us investigating AD in China as being worth studying. We also confirmed a point we had foreseen in previous stages of our research:

that users mainly have access to AD via movies. Yet, they claimed they would like to have it offered on TV as well. This makes sense taking into account that it is not very convenient for them to leave their homes on a regular basis or unaccompanied. Chinese AD users are satisfied with both the quantity and the quality offered to date, but this satisfaction could be the result of them still being oblivious to the wide scope of applications AD can have and their unawareness of their legal rights. Another explanation could be that our respondents were embarrassed to say otherwise for fear of not seeming grateful for what volunteers do for them. In fact, in the reception study carried out later, some users confessed they did not like having to refrain themselves from saying what they actually thought or wanted to be offered. As a result of this, they were willing to pay for this access service. Nonetheless, we know that, given the huge socioeconomic differences in China, not everybody –especially people with sight loss in the Chinese countryside– would or could pay for it.

Our participants stating that they frequently consume AD needs to be contextualised: they refer to doing so whenever AD is available, which varies according to the city. They mainly consume AD for entertainment purposes, with historical movies ranking among their favourite genres. This, though, has been found to be related to the users' age. Our participants would like to watch more foreign films as well and, although their wishes need to guide both academic research and practical action, we need to remember that, given how new AD is in China, meeting users' requirements cannot be divorced from overcoming several difficulties that even touch on legal issues (for example, the problem of copyright).

With regards to the fourth specific objective, three sets of unpublished in-house AD guidelines corresponding to different AD providers in Shanghai, Beijing and Guangzhou were analysed. Despite being in the capital and claiming to be the national providers of AD for the rest of the country, the guidelines provided by the China Braille Library are very brief and vague. Although the guidelines used in the Sun Yat-sen Library in Guangzhou were not drafted along with AD users, they encourage audio describers to include users in the revision of the script process, consult them when choosing what materials to audio describe and incorporate their feedback whenever they receive it. These guidelines also make an interesting contribution, which, to the best of our knowledge, has never been suggested before in any AD guidelines: they recommend the description of elements based on a tactile perspective (e.g. thin as a sheet of paper). The last set of

guidelines correspond to the most developed AD centre based in Shanghai: Sound of Light. Not only are their guidelines more complete, but they are also the most professional and specific. Given the fact that an AD user runs this AD production centre, the decisive role users have is underlined and proof of that is them taking part in all the AD production stages. All the AD guidelines under study shared the same main objective: training. Two of the three guidelines have been influenced by Western materials, more specifically the Netflix AD guidelines (Shanghai) and the ADLAB Audio Description Guidelines (Guangzhou). All our interviewees highlighted objectivity as one of the main lessons learnt from abroad.

The second general objective of this dissertation was to compare the Chinese case to that of Europe with regards to some specific issues. In order to achieve this aim, three specific objectives were set:

- a) To compare the Chinese audio describers' profile with that of the Europeans'.
- b) To compare the Chinese AD users' habits, needs and satisfaction with the European AD users.
- c) To compare the Chinese guidelines with the European ones.

Regarding the first specific objective, the results gathered in China were compared to those obtained in the ADLAB PRO project, the questionnaire of which was adapted for our research. Some relevant differences were observed, such as European audio describers associating AD with AVT whereas Chinese audio describers link it with Media Studies/Journalism and Linguistics (depending on whether the audio describer writes the AD script or voices it). Also, European audio describers have more experience in AD both in terms of years of experience and hours of material produced than our Chinese respondents, which makes sense taking into account that AD started in Europe earlier. Moreover, new AD production strategies were observed in the case of ADLAB PRO, such as AD translation, either carried out by humans or by using machine translation and post-editing. Although this is not a widespread practice yet, it shows that in Europe the research (or interest) in AD has resulted in ideas to speed up its production. It must be noted that the on-going Researching Audio Description Project³⁰ led by the research group TransMedia Catalonia at UAB is currently investigating the possibility of translating AD scripts from Spanish into Chinese. Were the researchers to obtain

³⁰ See <http://pagines.uab.cat/rad/> (Last accessed 22 April 2020)

promising results, translation could finally be implemented as an alternative way of creating AD in the country. Furthermore, unlike the Chinese case, European audio describers are involved professionally in AD, although they do not think they are well-paid. In addition, most European audio describers have received specific AD training, which is often related to AD applied to films and TV. In contrast, China offers more simple training, if any. In Europe training results in a certificate, whilst in China no certificates are offered. Also, the gateway to AD is varied in the European context while in China it mainly depends on acquaintances. Finally, European audio describers collaborate less with AD users but are more aware of the existence of guidelines than Chinese audio describers.

Nonetheless, some similarities between the two cases were also spotted. First, AD is a new discipline in both Europe and China. Second, audio describers are not usually asked for a certificate when being assigned an AD job. Third, audio describers agree on, for example, the little economic compensation received for creating AD is and on the social impact this access service has on society. Fourth, in both cases audio describers confirmed their interest in self-improvement and they usually do so by accumulating experience in the field.

As for the second specific objective, the results gathered in China were also compared to the ADLAB PRO results, although the questionnaire was highly modified (and shortened) to fit our research needs and users' background. Unlike our Chinese respondents, the ADLAB PRO respondents admitted having access to more kinds of AD, although most use it to watch films or TV, or when visiting a museum. It was observed that China is lagging behind because AD is still not even offered on TV, which is what users express a wish for the most and is probably the most popular medium for AD in many countries such as Poland (Jankowska and Walczak, 2019).

Regarding end-user satisfaction, results showed that most users were dissatisfied or very dissatisfied with the quantity of AD provided in Europe, although they were quite satisfied with the quality of the products that were offered. We also observed Chinese users being more satisfied in terms of quality, but they also claimed to be quite satisfied with the quantity provided, probably because they are still not aware of all the applications audio description can have.

Regarding the third specific objective, Europe has a much longer-standing tradition when it comes to creating guidelines for AD and even standards, although many of them are intuitive and based on personal experiences. There are now, though, new guidelines that break with this tradition, such as the ADLAB guidelines, which, for example, provide a tool to make an informed selection of what elements to include in the AD. Also, they put an end to the perennial dispute over whether it is objective or subjective by suggesting that AD is always the latter for it is invariably based on the interpretation of the describer (Remael, Reviere & Vercauteren, 2015, p. 16).

Our analysis of Chinese guidelines and the comparison with the European ones leads us to believe that there are many similarities between the Chinese and the European “traditional” AD guidelines. First, they have been drafted with no empirical base, but depending only on personal experience. Second, they all include, to a greater or lesser extent, what to describe, the AD style and when to describe. Third, the main objective in all cases has always been to train audio describers. Yet, a clear difference between the European and the Chinese case has been identified: the Chinese guidelines focus on being objective while this concept seems to already be outdated in Europe. In fact, we believe Chinese guidelines resemble Europe’s more traditional guidelines. This makes sense taking into account that China is behind western countries (and even some Asian ones such as Japan) in terms of AD development.

The third main objective is aimed at determining whether text-to-speech AD can be accepted in movies in China. In order to accomplish this last objective, the following specific objectives were defined:

- a) To evaluate TTS AD in Chinese by comparing it with standard human-voiced AD on key features.
- b) To elucidate whether TTS AD could be accepted as either an interim or a permanent solution, or both, in the Chinese context.

We first checked the comprehension of the clips screened to verify objectively that users had understood what both the TTS and the human voice had described. The average of correct answers was, in our opinion, quite low given the length of the clip and the difficulty of the questions proposed. No statistical differences were observed when we compared the results obtained in the case of the clip audio described with TTS and those gathered after watching the human-voiced clip. Thus, the limited comprehension does not

seem to be related to the type of voice used. Given that younger audiences did not perform statistically better in terms of comprehension, we are inclined to believe that the way AD is currently performed needs to be revised. In fact, the users (26.9%) answering the questionnaire designed by Sound of Light already indicated that they sometimes did not understand some parts of the AD. Although this falls outside the scope of our study, we would venture to suggest that it could be due to, for example, the cognitive load imposed by the AD or to their educational background.

In our reception study we compared a synthetic voice with a human voice. A selection of some key features related to the quality of the voices was made and we found out that, in general, human voices were more highly rated. This allows us to say that the quality of the human voice used in the study is still not comparable to that of the TTS. This finding is in line with the majority of Chinese users in our sample preferring human voices when watching an audio described movie. Yet, they accepted TTS AD both as an interim (97.5%) and permanent solution (62.5%). Our numbers are similar (and even more positive in the case of the possibility of having TTS AD as an interim solution) to those obtained in previous studies, such as those by Fernández-Torné and Matamala (2015) and the Polish team testing TTS AD (Walczak, 2010; Szarkowska, 2011; Drożdż-Kubik, 2011; Mączyńska & Szarkowska, 2011; Szarkowska & Jankowska, 2012). Although the quality of the voices tested might be different depending on the country, users have proved to be optimistic about this type of delivery, probably because in many cases they prefer having some sort of AD rather than not having it. We believe that, given the state of AD in China, Chinese AD users still give priority to quantity rather than quality. That is to say that any option that means more access to audio-visual content is very welcome. Yet, the potential of TTS AD cannot be denied in the light of some users not even noticing that one of the clips was being voiced by a synthetic voice. Nonetheless, TTS AD should be ideally offered through, for example, earphones so that users could consume it while watching a movie with sighted viewers. In fact, some users expressed this idea more than once during our research. In December 2019, before finishing this dissertation, we came to know that from January 2020 on, 50 cinemas in Shanghai will be offering AD through earphones thanks to a project called “Most beloved cinemas: Accessible film-watching” (*Zhi'ai yingyuan: wuzhang'ai guanying* 至爱影院——无障碍观影).³¹

³¹ See <http://www.crifst.ac.cn/index.php?c=show&id=1293> (Last accessed 8 December 2019)

Despite us not making a previous selection of various voices that could be used in the reception study, the synthetic voice chosen has proven to be accepted in most cases. In fact, 20% of our participants preferred it over the human voice and almost 30% did not have a clear preference. It is expected that the results yielded could be even be more promising by using the latest versions of commercial TTS. Our study also shows that intonation, and, by extension, prosody, is still a pending subject for TTS in Chinese. Yet, this issue does not seem to have an extremely obvious impact on the acceptance of TTS AD, although it was a much-repeated comment. Strangely enough, no statistical proof can be provided to prove, once again, that frequency of use of TTS is related to its acceptance. This result could be explained by our sample size being too small. The participants of our study also made some comments that lead us to believe that the AD currently offered in China needs to be revised. For example, they complained about the choice of human voices for them not always being different enough from those of the film characters.

Despite the extra challenges Chinese TTS has to face, our results open the doors to this form of AD delivery that not only expedites the production of AD, but also allows for the expansion of AD boundaries in the country by making it less volunteer-dependent. According to the results TTS AD has obtained in our reception study, we believe it is a feasible and plausible solution for China until AD is professionalised and more monetary and human resources are invested or socio-politically prioritised. Also, and since TTS is constantly improving, it is reasonable to think that the acceptance of TTS AD will only increase over time. To illustrate the improvements TTS is undergoing, one might mention the Chinese documentary “China Reinvents Itself” (Liu & Xu, 2018). According to Gambier and Jin (2018), it is claimed to be the first documentary in the world using artificial voices. The voice was cloned from the well-known voice artist Li Yi. The quality of the voice was so superb that even acquaintances of Li Yi could not tell the difference between the TTS and Li Yi’s real voice. Another example of the growing interest in synthetic voices in China are the avatars launched by the company Sogou, which are used as newsreaders by the government’s Xinhua news agency. The Chinese-native web crawler Baidu has also launched a system called Deep Voice that learns how to imitate a person’s voice using just three seconds of voice sample data (Arik *et al.*, 2018). This system has been conceived as a way to personalise virtual assistants such as Apple’s Siri or “Baidu’s Mandarin virtual assistant platform DuerOS, which supports 50 million

devices in China with human-machine conversational interfaces” (Peng, 2018). Voice cloning has also been found to be useful in healthcare for helping patients who have lost their voices by building a duplicate, as well as for “the entertainment industry and in social media as a tool for satirists” (Peng, 2018). Although no mention has been found about its possible use in the AD field, we believe these TTS-related improvements will definitely help to push AD and make it gain more visibility. Notwithstanding, the advances TTS are experimenting with are more than likely to raise concerns and bring up new issues that will have to be taken care of:

AI could potentially downgrade voice identity in real life or with security systems. For example, voice technology could be used maliciously against a public figure by creating false statements in their voice. A BBC reporter’s test with his twin brother also demonstrated the capacity for voice mimicking to fool voiceprint security systems. (Peng, 2018)

These concerns, though, fall outside the scope of this thesis and, therefore, are left for future scholars working on this exciting avenue of research. Nonetheless, this exciting context of TTS improvements and advances justifies the pertinence and innovativeness of this thesis, which modestly contributes to developing a non-commercial field that has been widely ignored by Chinese academia.

6.3. Future lines of research

This PhD thesis has opened new avenues of research that are worth exploring. This section will describe some of them, excluding the ones already mentioned in the conclusions section in each article.

To begin with, cities such as Beijing and Guangzhou should be further explored. As previously mentioned, this is due to the author spending most of her research stays in the city of Shanghai. It would be interesting to research the Bright Cinema Project (光明影院) in Beijing, which was launched by the Communication University of China in 2017. Said project has produced around 200 movies in the past years and plans to increase its production to two audio described movies per week.

Also, since one of the main challenges the development of AD has to face is copyright, it would be advisable to investigate how copyright is dealt with in other countries so that Chinese AD providers can follow other countries’ examples or, at least, come up with

some ideas. Jankowska and Fidyka (2017) already started looking at current practices regarding how copyright was handled mostly in Europe and how to create AD legally. Something similar could be explored for the Chinese case to help Chinese legislators make informed decisions on how to modify the existing Copyright Law of the People's Republic of China, which does not mention AD at all.

Furthermore, more reception studies should be carried out to draft experiment-based AD guidelines or even standards. Also, the AD guidelines consulted do not represent the whole country and more local guidelines should be found to obtain a more representative picture of how AD is currently performed. This could be a preliminary step before drafting unified guidelines based on empirical research for them to be used as a national standard all over China. In order to achieve this, Leung's (2018) guidelines could be taken as a starting point. Moreover, and since the AD sessions in Shanghai are mainly attended by elderly users (except for those especially targeted at children with sight loss in specific schools), the concept of age-oriented AD merits research.³² The underlying reason for this division is that age differences in word and language processing affect the viewing process as a critical cognitive one, especially in the case of a morphosyllabic writing such as Chinese. This has started to be researched by Aleksandrova (2019). In addition, access services compatible with AD, such as AI and AST, should be introduced to improve the service currently offered and tested in combination with AD to assess their reception.

Moreover, crowdsourcing applied to AD could be studied in China. This has already been explored by Jankowska (2018) as part of the ADDit! project. Jankowska (2018) organised the creation of two ADs for two films for a music film festival in Krakow by means of crowdsourcing. The outcome was positive and it was concluded that, for example, the quality of the AD could be improved thanks to brainstorming together and audio describers having more time to audio describe their part. Yet, some drawbacks were also underlined, such as the high dependence on careful planning and all participants meeting the deadlines. Researchers interested in producing AD by this means to expedite the production of AD in China could consult the conclusions reached and improvements suggested by Jankowska (2018) and try this technique in the country.

³² The Easit Project (<http://pagines.uab.cat/easit/en>, last accessed 5 March 2020) led by TransMedia Catalonia at the Autonomous University of Barcelona is currently investigating how to apply easy-to-read practices to audiovisual content.

In addition, and since we have proven that human voice is still the preferred solution over plain synthetic voices (non-enacted), new ways to improve the delivery of human-voices AD can be explored. This is important because some users observed that using a human voice was not enough to offer a quality service. For example, AD and enacted AD, both in human voice, are yet to be compared and further studied. As Sahasrabuddhe (2019) suggested, in order to convey emotions which are elicited in sighted people by film visuals, verbal enactment of AD script can be a method worth researching. This researcher proved with statistical data that enacted AD improves the film watching experience as well as level of enjoyment and level of understanding for viewers with sight loss.

Also, it would be interesting to follow Perego's (2019) advice, which stresses the importance of involving people with no visual loss as well to determine whether they can benefit from, not just tolerate, AD. Perego (2019) believes that the paucity of this type of studies hinders this accessibility service from vindicating a wider presence and scope. Therefore, including people with no sight loss could result in wider availability of AD in China in the future.

In terms of methodology, all these ideas could be studied using more innovative techniques, i.e., taking advantage of recent technological developments such as eye-tracking, electroencephalography and other psychophysiological measures. As Díaz-Cintas and Szarkowska (2020, p. 5) point out, studying complex semiotic composites as heterogeneous and multimodal as audio-visual productions necessitates the use of sophisticated tools.

Finally, it is strongly suggested to research accessible filmmaking to reduce or even eliminate the need for access services implemented in an *ex-post* stage. For example, the Enhancing Audio Description Project, led by the University of York, has investigated three audio-related techniques to reduce the number of verbal descriptions needed. Chinese researchers could also explore how to make their media more accessible by incorporating some techniques during the creation process. The accessible filmmaking guide created in a project led by the research group GALMA³³ could also be of help, since it is intended for filmmakers and other professionals within the film industry who wish to become accessible filmmakers.

³³ See <http://galmaobservatory.webs.uvigo.es/projects/accessible-filmmaking-guide/> (Last accessed 22 April 2020)

6.4. Limitations

As with any other academic research, the findings of this PhD thesis have to be seen in light of some limitations, the main ones being the small size of our samples to yield statistically powerful results (due to the limited access to AD users, which seems to be common worldwide, as Orero *et al.* [2018]) point out) and the results mostly informing us about Shanghai's AD situation and other major cities such as Beijing and Guangzhou, leaving the rest of Mainland China unresearched. This is the reason why our dissertation should be taken as a preliminary glimpse into both the current state of AD in Mainland China and the acceptance of TTS AD. Yet, Orero *et al.* (2018) suggest that AVT researchers should recruit at least 30 participants with the expectation that around 5 will not provide complete data. Therefore, our sample sizes are all above the minimum expected in the AVT field.

Since our reception study is a replication of previous studies testing TTS AD, we can proudly claim that we have responded to the call of Di Giovanni (2018) and Olalla-Soler (2019), for example, to produce more replication studies by contributing to a cross-national and cross-linguistic reach that few AVT reception studies have. This type of collaborative studies, albeit not attractive (Szarkowska, 2018), reveal whether findings are consistent and can capture the heterogeneity that characterises real life. Nonetheless, in the case of our reception study it must be borne in mind that:

While not as rigorous as experiments, quasi-experiments are often a necessary alternative in AVT that allow researchers to assign participants to groups based on characteristics and the factors mentioned above. Quasi-experiments, however, can lack internal validity and run a higher risk of having a more limited, if any, generalisability and replicability. (Orero *et al.*, 2018, p. 108)

The huge differences among Chinese provinces, and even within one same province, have already been presented in the introductory chapter of this thesis and have been mentioned in the articles. Therefore, they will not be developed in this section again. Instead, we will focus on the two limitations that have not been specified yet. The first concerns the very little or no prior research on AD in Mainland China and Leung's (2018) thesis being released right after our surveys were circulated and the reception study was performed. Having the chance to read and learn from a previous and recent reception study carried out with Greater China users would have been very helpful to solve questions and have

more materials to make informed decisions. Obviously, the time constraints our thesis was subject to are responsible for this, especially taking into account that the journal's acceptance process and time needed to get a response from the editor had to be borne in mind at all times. In fact, shaping a PhD thesis as a compendium of articles has another disadvantage as well, for example, not being able to make modifications to already accepted (and even published) papers. This is also considered to be a limitation because the writing of a PhD dissertation is a long process that requires novice researchers to constantly revise their work to reach mature enough conclusions.

The second limitation is linked to the conflicts arising from cultural bias and other personal issues, which cannot be fully avoided but it is undeniable that they condition the research design and results interpretation. Undoubtedly, researchers are biased due to their cultural background for certain phenomena. In our case, we have tried to make sure that all the materials suited the Chinese context, but the researcher, despite speaking fluent Chinese, is not a native speaker and it is inevitable that this fact had some consequences that could not be detected at the time. Nonetheless, being a foreign researcher in China has had some advantages. For example, many users and audio describers were curious about us being interested in AD in China. They warmly welcomed us, showed great interest in our research and were always happy to help. The lack of general interest in this field also helped us in the sense that AD stakeholders are not used to being at the centre of academic attention and found it interesting and amusing to participate in a project that aimed at improving the provision of the service. Furthermore, the fact that we speak Chinese allowed us to avoid intermediaries and to have direct contact with our target AD stakeholders and, therefore, to obtain first-hand data, which otherwise would have been impossible to have access to.

6.5. Knowledge transfer to society

The author of this thesis places great importance on the idea of transferring the knowledge acquired throughout the years devoted to her PhD thesis to the Chinese society. This was deemed necessary because MA is still both an unresearched topic by Chinese academia and an untaught subject at Chinese universities. Should China be interested in keeping up with AVT international research trends, it should start paying more attention to this

emerging field that is consolidating rapidly. Due to the importance attached to knowledge transfer in our sociocultural context, we decided to devote a specific subsection in this concluding chapter to listing chronologically the activities the author took part in, which range from workshops and academic articles to interviews. Our aim was to both disseminate results and raise awareness about the importance of researching MA in China:

- a) The author offered a lecture (“History and current situation of audio description in Spain”) at the “Sound of Light Accessible Movies” Workshop organised by Sound of Light (Shanghai, 21 October 2017).
- b) The author offered a two-day AD workshop at the Shanghai International Studies University (April 2019). During the first session the students were shown the ropes of MA with special focus on AD and in the following session Prof. Han Ying, an AD user running the AD production centre Sound of Light, was invited to comment on the ADs the students had created for homework.
- c) The author was interviewed to present her research on AD in China at the online media platform ShanghaiEye (21 May 2019). It can be accessed at <https://www.shanghaieye.com.cn/an-expats-push-for-a-more-accessible-shanghai/>.
- d) The author offered a two-hour workshop at the Hangzhou International Studies University (21 June 2019) to introduce MA, with special focus on AD, to first-year students studying a BA in Translation.
- e) The author was interviewed to present her PhD research for the Confucius Institute’s WeChat official account (see Annex 6.1).
- f) The author was interviewed to present her PhD research for the Shanghai International Studies University’s WeChat official account (see Annex 6.2.).
- g) The author offered a presentation (“Audio description in Europe”) at an AD live session organized by the China Global Philanthropy Institute at the Lyceum cinema (Shanghai, 9 August 2019).
- h) The author wrote an academic article along with her Chinese advisor entitled “Shiting fanyi xin fazhan: yuyin hecheng zai koushu yingxiang zhong de yingyong. Jiyu zhendui shizhang renshi de jieshou shiyan yu diaocha” (视听翻译新发展：语音合成在口述影像中的应用——基于针对视障人士的接受实验与调查, New developments in audio-visual translation: The use of synthetic voices in audio description. An acceptance experiment for people with sight loss)

and was accepted by the academic journal *Dongfang Fanyi* (东方翻译, East Journal of Translation) to disseminate our final results so that more Chinese stakeholders could have access to our research. This was decided after realising that not enough Chinese individuals interested in AD could have access to our results if they were written in English.

We would also like to point out a couple of issues that make us particularly proud. First, thanks to our dissemination activities in China, a Chinese PhD candidate is now working on a thesis revolving around AD at the Shanghai International Studies University, which is temporarily entitled as “A multimodal approach to interpreting pedagogy: researching the use of audio description in initiation practice”. This will be the first thesis carried out by a Chinese citizen from the Mainland focusing on AD —and, by extension, on MA— and we hope it will be the first of many academic studies that will lay the foundations for MA and awake the academic interest in this discipline in China. Second, after the workshop offered at the Shanghai International Studies University, the lecturer in charge of the subject (Dr Xiao Weiqing) decided to repeat the experience and invite an AD user every year to include MA, especially AD, in the subject’s curricula. Third, since we carried out the reception study in China, Sound of Light has continued investigating the idea of TTS AD and is now formally collaborating with Ke Da Yuyin (科大语音), the company whose synthetic voice was used for our experiment. This company is now sharing more natural voices and they are currently audio describing a full movie. The continuation and real application of our research makes us very happy and we believe we have truly contributed to pushing AD forward in China, and therefore have some social impact on the country. This is in line with the AR approach we had taken from the very beginning: “No action without research, no research without action” (Lewin, 1946), which we believe best suits research that can have an immediate social application. Our reception study has resulted in a change in a community and we hope more research is carried out in this direction so that more feedback is collected to keep on improving the offer of TTS AD.

Chapter 7. Bibliography

- Adelman, C. (1993). Kurt Lewin and the Origins of Action Research. *Educational Action Research, 1*(1), 7–24.
- ADLAB PRO. (2017). *Audio Description Professional: Profile Definition*. Retrieved from: <https://www.adlabpro.eu/wp-content/uploads/2018/04/IO2-REPORT-Final.pdf>
- Aguado Díaz, A. L. (1995). *Historia de las deficiencias*. Madrid: Escuela Libre Editorial, Fundación ONCE, 40–41.
- Agulló, B., Matamala, A., & Orero, P. (2018). From Disabilities to Capabilities. Testing Subtitles in Immersive Environments with End Users. *HIKMA. Translation Studies Journal, 17*, 195–220.
- Aleksandrova, E. (2019). Age-oriented Audio Description. Paper presented at Media for All 8. Stockholm, 19 June 2019.
- Arik, S. Ö., Chen, J., Peng K., Ping, W., & Zhou, Y. (2018). Neural Voice Cloning with a Few Samples. In S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, & N. Cesa-Bianchi (Eds.), *Proceedings of the 32nd Conference on Neural Information Processing Systems (NIPS 2018)* (pp. 10040-10050). Montréal, Canada. Retrieved from: <https://arxiv.org/pdf/1802.06006.pdf>
- Avery, K. (2016, 7 March). Disability and the Three Traditional Chinese Belief Systems: Buddhism, Taoism and Confucianism. *China Source*. Retrieved from: <https://www.chinasource.org/resource-library/articles/disability-and-the-three-traditional-chinese-belief-systems>
- Bal, M. (1997). *Narratology: An Introduction to the Theory of Narrative*. Toronto: University of Toronto Press.
- Benecke, B. (2004). Audio Description. *Meta, 49*(1), 78–80.
- Benecke, B., & Dosch, E. (1997). *Wenn aus Bildern Worte Werden. Ein Handbuch für Filmbeschreiber*. München: Bayerischer Rundfunk.
- Benecke, B., & Dosch, E. (2004). *Wenn aus Bildern Worte werden. Durch Audio-Description zum Hörfilm*. München: Bayerischer Rundfunk.
- Bourne, J., & Jiménez Hurtado, C. (2007). From the Visual to the Verbal in Two Languages: A Contrastive Analysis of the Audio Description of *The Hours* in English and Spanish. In J. Díaz-Cintas, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 175–187). Amsterdam: Rodopi.

- Branje, C., & Fels, D. (2012). LiveDescribe: Can Amateur Describers Create High-Quality Audio Description? *Journal of Visual Impairment and Blindness*, 106(3), 154–165. doi: 10.1177/0145482X1210600304
- Braun, S. (2008). Audiodescription Research. State of the Art and Beyond. *Translation Studies in the New Millennium*, 6, 14–30.
- 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. 227–238). Amsterdam/New York: Rodopi.
- Cabeza-Cáceres, C. (2013). *Audiodescripció i Recepció. Efecte de la Velocitat de Narració, l'Entonació i l'Explicitació en la Comprensió Fílmica* (PhD dissertation). Universitat Autònoma de Barcelona, Spain.
- Campbell, A., & Uren, M. (2011). “The Invisibles”... Disability in China in the 21st century. *International Journal of Special Education*, 26(1), 12–24.
- Casas-Tost, H., & Rovira-Esteva, S. (2018). A Quest for Effective and Inclusive Design of Chinese Characters in Subtitling. *International Journal of Asian Language Processing*, 28(1), 31–47.
- Chao, Y. (赵雅丽) (2002). *Yuyan Shijie Zhong de Liudong Guangying. Koushu Yingxiang de Lilun Jiangou* (语言世界中的流动光影。口述影像的理论建构) [The moving lights in the world of language. Theory of audio description]. Taipei: Wunan Publishing Co.
- Chesterman, A. (2007). Bridge Concepts in Translation Sociology. In M. Wolf, & A. Fukari (Eds.), *Constructing a Sociology of Translation* (pp. 171–183). Amsterdam/Philadelphia: John Benjamins.
- Chiaro, D. (2007). The Effect of Translation on Humour Response: The Case of Dubbed Comedy in Italy. In Y. Gambier, M. Shlesinger, & R. Stolze (Eds.), *Doubts and Directions in Translation Studies: Selected Contributions from the EST Congress, Lisbon 2004* (pp. 137–152). Amsterdam/Philadelphia: John Benjamins.
- China Disabled Persons’ Federation. (2007). *2006 Nian Di-Er Ci Quanguo Canjiren Chouyang Diaocha Zhuyao Shuju Gongbao* (2006 年第二次全国残疾人抽样调查主要数据公报) [Report on the results of the second national sample survey on persons with disabilities]. Retrieved from: <http://www.cdpcf.org.cn/sjzx/cjrgk/200711/t20071121_387540_1.shtml>

- China Disabled Persons' Federation. (2016). *Law on the protection of persons' with disabilities*. Retrieved from: http://www.cdpcf.org.cn/english/Resources/lawsregulations/201603/t20160303_542879.shtml
- Chmiel, A., & Mazur, I. (2012). AD Reception Research: Some Methodological Considerations. In E. Perego (Ed.), *Emerging Topics in Translation: Audio description* (pp. 57–74). Trieste: EUT.
- Chmiel, A., & Mazur, I. (2016). Researching Preferences of Audio Description Users. Limitations and Solutions. *Across Languages and Cultures*, 17(2), 271–288.
- Cryer, H., & Home, S. (2008). *Exploring the Use of Synthetic Speech by Blind and Partially Sighted People. Literature review #2*. Birmingham: RNIB Centre for Accessible Information (CAI). Retrieved from: https://www.rnib.org.uk/sites/default/files/2008_11_Exploring_use_of_synthetic_speech.doc
- Cryer, H., & Home, S. (2009). *User Attitudes Towards Synthetic Speech for Talking Books*. Birmingham: RNIB Centre for Accessible Information (CAI). Retrieved from: https://www.rnib.org.uk/sites/default/files/2009_05_User_attitudes_synthetic_talking_books.doc
- Cryer, H., & Home, S. (2010). *Review of Methods for Evaluating Synthetic Speech. Technical report #8*. Birmingham: RNIB Centre for Accessible Information (CAI). Retrieved from: https://www.rnib.org.uk/sites/default/files/2010_02_Evaluating_synthetic_speech_review.doc
- Dauncey, S. (2007). Screening Disability in the PRC: The Politics of Looking Good. *China Information*, 21(3), 481–506. doi: 10.1177/0920203X07083324
- Deloitte (n.d.). “China's Film Industry – a New Era”. Retrieved from <https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/technology-media-telecommunications/deloitte-cn-tmt-china-film-industry-en-161223.pdf>
- Derbring, S., Ljunglöf, P., & Olsson, M. (2009). TS: Light-weight Automatic Reading of Subtitles. In K. Jokinen, & E. Bick (Eds.), *NODALIDA 2009 Conference Proceedings* (pp. 272–274).
- Díaz-Cintas, J. (2005). Audiovisual Translation Today – A Question of Accessibility for All. *Translation Today*, 4, 3–5.

- Díaz-Cintas, J., & Szarkowska, A. (2020). Introduction: Experimental Research in Audiovisual Translation – Cognition, Reception, Production. *JoSTrans: The Journal of Specialised Translation*, 33, 3–16.
- Dick, B. (1993). “You want to do an action research thesis? -- How to conduct and report action research (including a beginner's guide to the literature)”. *Action Research & Action Learning for Community & Organisational Change*. Retrieved from http://www.aral.com.au/resources/arthesis.html#a_art_methchap
- Di Giovanni, E. (2012). Italians and Television. A Comparative Study of the Reception of Subtitling and Voice-over. In S. Bruti, & E. Di Giovanni (Eds.), *Audiovisual Translation across Europe. An Ever-changing Landscape* (pp. 171-190). Bern: Peter Lang.
- Di Giovanni, E. (2018a). Audio Description for Live Performances and Audience Participation. *JoSTrans: The Journal of Specialised Translation*, 29, 189–211.
- Di Giovanni, E. (2018b). Audio Description and Reception-centred Research. In E. Di Giovanni, & Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation* (pp. 225–250). Amsterdam/Philadelphia: John Benjamins.
- Di Giovanni, E., & Romero-Fresco, P. (2019). Are We All Together across Languages? An Eye Tracking Study of Original and Dubbed Films. In I. Ranzato, & S. Zanotti (Eds.), *Reassessing Dubbing. Historical Approaches and Current Trends* (pp. 126-144). Amsterdam: John Benjamins.
- Drożdż-Kubik, J. (2011). *Harry Potter i Kamień Filozoficzny słowem malowany – czyli badanie odbioru filmu z audiodeskrypcją z syntezą mowy* (Masters dissertation). Jagiellonian University, Poland.
- Efe Agro (2018). “La población rural baja un 9 % desde el año 2000 y solo crece en Madrid y Canarias”. Retrieved from <https://www.efeagro.com/noticia/lpoblacion-rural-baja-9-desde-ano-2000/>
- Fei, X. (2011). *Audio Description in the Chinese Community* (Masters dissertation). Capital University of Economics and Business, Canada.
- Feng, J. (冯纪元) (2018). Koushu yingxiang de banquan guizhi tanjiu: yi yingshi zuopin wei li (口述影像的版权规制探究——以影视作品为例) [Study on the copyright of AD: Taking audiovisual works as an example]. *Henan Keji* (河南科技) [Henan Science and Technology], 3, 36–38.

- Fernández-Torné, A. (2016). *Audio Description and Technologies. Study on the Semi-automatisation of the Translation and Voicing of Audio Descriptions*. PhD thesis. Autonomous University of Barcelona, Spain.
- Fidyka, A., & Matamala, A. (2018). Audio Description in 360 Videos: Results from a Focus Group in Barcelona. *Translation Spaces*, 7(2), 285–303.
- Flis, G., Sikorski, A., & Szarkowska, A. (2020). Does the Dubbing Effect Apply to Voice-over? A Conceptual Replication Study on Visual Attention and Immersion. *JoSTrans: The Journal of Specialised Translation*, 33, 41–69.
- Fresno, 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*. Universitat Autònoma de Barcelona, Spain.
- Fresno, N., Castellà, J., & Soler Vilageliu, O. (2016). ‘What Should I Say?’ Tentative Criteria to Prioritize Information in the Audio Description of Film Characters. In A. Matamala, & P. Orero (Eds.), *Researching Audio Description. New Approaches* (pp. 143-167). New York: Palgrave Macmillan.
- Fryer, L. (2016). *An Introduction to Audio Description*. London: Routledge. doi: 10.4324/9781315707228
- Fryer, L. (2018). The Independent Audio Describer is Dead: Long Live Audio Description! *Journal of Audiovisual Translation*, 1(1), 170–186.
- Fryer, L., & Walczak, A. (2017). Creative Description: The Impact of Audio Description Style on Presence in Visually Impaired Audiences. *British Journal of Visual Impairment*, 35(1), 6–17.
- Gambier, Y. (2003). Introduction. *The Translator*, 9(2), 171–189. doi: 10.1080/13556509.2003.10799152
- Gambier, Y. (2018). Translation Studies, Audiovisual Translation and Reception. In E. Di Giovanni, & Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation* (pp. 43-66). Amsterdam: John Benjamins.
- Gambier, Y., & Jin, H. (2018). Audiovisual Translation in China: A Dialogue between Yves Gambier and Haina Jin. *Journal of Audiovisual Translation*, 1(1), 26–39.
- Gambier, Y., & Jin, H. (2019). A Connected History of Audiovisual Translation: Elements for Consideration. *Translation Spaces*, 8(2), 193–230.
- Gerber-Morón, O. (2018). *Subtitle Segmentation Quality across Screens* (PhD dissertation). Universitat Autònoma de Barcelona, Spain.

- 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. 11-33). London: Palgrave MacMillan.
- Greco, G.M. (2019). Accessibility Studies. Abuses, Misuses and the Method of Poietic Design. In C. Stephanidis (Ed.), *HCI International 2019. Late Breaking Papers* (pp. 15–27). Cham: Springer.
- Greco, G. M., & Jankowska, A. (forthcoming), Media Accessibility within and Beyond Translation Studies. In Ł. Bogucki, & M. Deckert (Eds.), *Handbook of Audiovisual Translation and Media Accessibility*. Cham: Palgrave Macmillan.
- Greening, J., & Rolph, D. (2007). Accessibility: Raising Awareness of Audio Description in the UK. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 127–138). Amsterdam/New York: Rodopi.
- Greening, J., Petré, L., & Rai, S. (2010). A Comparative Study of Audio Description Guidelines Prevalent in Different Countries. Retrieved from: http://audiodescription.co.uk/uploads/general/RNIB._AD_standards.pdf
- Hallett, S. (2005, 6 October). “One eye on China: Challenging disability”. *Ouch!* Retrieved from: <http://www.bbc.co.uk/ouch/features/one-eye-on-china-challenging.shtml>
- Herman, D. (2009). *Basic Elements of Narrative*. Oxford: Wiley-Blackwell.
- Herman, L., & Vervaeck, B. (2005). *Handbook of Narrative Analysis*. Lincoln - London: University of Nebraska Press.
- Hermosa, Irene (forthcoming). Delivery Approaches in Audio Description for the Scenic Arts. To appear in *Paralleles*.
- Hernández-Bartolomé, G., & Mendiluce-Cabrera, A.I. (2009). How can images be translated? Audio description, a challenging audio-visual and social gap-filler. *Hermeneus, Revista de Traducción e Interpretación*, 11, 1–17
- Hinterleitner, F., Neitzel, G., Möller, S., & Norrenbrock, C. (2011). An Evaluation Protocol for the Subjective Assessment of Text-to-speech in Audiobook Reading Tasks. In *Proceedings of the Blizzard Challenge Workshop* (n. p.). International Speech Communication Association (ISCA).
- Hjelmquist, E., Jansson, B., & Torell, G. (1990). Computer-oriented Technology for Blind Readers. *Journal of Visual Impairment and Blindness*, 17, 210–215.

- Iaia, P.L. (2015). *The Dubbing Translation of Humorous Audiovisual Texts*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Igareda, P. (2011). The Audio Description of Emotions and Gestures in Spanish-spoken Films. In A. Serbal, A. Matamala, & J.L. Savaur (Eds.), *Audio-visual Translation in Close-up* (pp. 223–238). Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien: Peter Lang.
- Iglesias Fernández, E. 2010. La dimensión paralingüística de la audiodescripción: un acercamiento multidisciplinar. In C. Jiménez Hurtado, C. Seibel, & A. Rodríguez (Eds.). *Un corpus de cine: Fundamentos teóricos y aplicados de la audiodescripción* (pp. 21–29). Granada: Tragacanto.
- Iglesias Fernández, E., Martínez Martínez, S., & Chica Núñez, A. (2011). Reception Studies in Audio Description and Interpreting Quality Assessment: A Case of Cross-Fertilization. Paper presented at the 4th Media for All Conference, 28 June-1 July, London.
- Ioannidis, J. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), 696–701. doi: 10.1371/journal.pmed.0020124
- ISO (International Organization for Standardization). (2015). TS 20071-21 Information Technology — User Interface Component Accessibility — Part 21: Guidance on Audio Description.
- ITU (International Telecommunication Union). (1994). Telephone Transmission Quality Subjective Opinion Tests. A Method for Subjective Performance Assessment of the Quality of Speech Voice Output Devices. Geneva, Switzerland.
- ITU (International Telecommunication Union). (2017). *Question 7/1: Access to Telecommunication/ICT Services by Persons with Disabilities and with Specific Needs*. Retrieved from: https://www.itu.int/dms_pub/itu-d/opb/stg/D-STG-SG01.07.4-2017-PDF-E.pdf
- Iturregui-Gallardo, G. (2019). *Audio subtitling: Strategies for Voicing Subtitles and Their Effect on Film Enjoyment and Emotion*. PhD thesis. Autonomous University of Barcelona, Spain.
- Jakobson, R. (2000 [1959]). On Linguistic Aspects of Translation. In L. Venuti (Ed.), *The Translation Studies Reader* (pp. 113-118). London and New York: Routledge.
- Jankowska, A. (2015). *Translating Audio Description Scripts: Translation As a New Strategy of Creating Audio Description*. Berna: Peter Lang.

- Jankowska, A. (2018). Translation, Crowdsourcing, Collaboration and Quality in Audio Description. Paper presented at the UMAQ Conference, 4 June 2018, Barcelona.
- Jankowska, A., & Walczak, A. (2019). Audio Description for Films in Poland: History, Present State and Future Prospects. *JoSTrans: The Journal of Specialised Translation*, 32, 236–261.
- Jankowska, A., & Fidyka, A. (2017). Audio Description Legally Speaking. Paper presented at I the Media Accessibility Platform International Conference, 5 October 2017.
- Jankowska, A., & Fryer, L. (2017). Translating Audio Description Scripts... into English. *SKASE Journal of Translation and Interpreting*, 10(2), 2–16.
- Jia, X., & Zhao, Y. (2008). *China Civil Society Report: An Overview of Social Work with the Disabled in China*. Nautilus Institute. Retrieved 3 from: <http://www.nautilus.org/index.html>
- Kobayashi, M., O'Connell, T., Gould, B., Takagi, H., & Asakawa, C. (2010). Are Synthesized Video Descriptions Acceptable? In *ASSETS '10: Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 163–170). New York: ACM.
- Koverienė, I., & D. Satkauskaitė (2018). Lithuanian Viewers' Attitude Towards Dubbed Animated Films. *Readings in Humanities*, 3, 67–84.
- Kruger, J-L. (2009). The Translation of Narrative Fiction: Impostulating the Narrative Origo. *Perspectives: Studies in Translatology*, 17(1), 15–32.
- Kruger, J-L. (2010). Audio Narration: Re-narrativising Film. *Perspectives: Studies in Translatology*, 18(3), 231–249.
- Lambert, A. (2016). The Usefulness of the Useless': Rethinking Disability Through Classical Daoist Thought. In D. Y. Schumm, & M. Stoltzfus (Eds.), *Disability and World Religions: An Introduction* (n.p.). Waco: Baylor University Press.
- Lavaur, J.M., & Bairstow, D. (2011). Languages on the Screen. Is Film Comprehension Related to the Viewers' Fluency Level and to the Language in the Subtitles. *International Journal of Psychology*, 46(6), 455-462.
- Leung, D. (2015). Audio Description in Hong Kong. In R. Baños, & J. Díaz-Cintas (Eds.), *Audiovisual Translation in a Global Context* (pp. 266–281). London: Palgrave Macmillan.

- Leung, D. (2018). *Audio Description of Audiovisual Programmes for the Visually Impaired in Hong Kong* (PhD dissertation). University College London, United Kingdom.
- Lewin, K. (1946). Action Research and Minority Problems. In G.W. Lewin (Ed.), *Resolving Social Conflicts. Selected Papers on Group Dynamic* (pp. 143-152). Washington DC: American Psychological Association.
- Lewis, C. (2014). *"Use of the Useless": Assessing Depictions of Disability in the Zhuangzi*. Bachelor's diss. Brigham Young University. Retrieved from: https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1017&context=libraryrg_studentpub
- Li, D. (李东晓) (2013). *Tingjian, Kanjian: Yingshi Meiti de Wuzhang'ai Chuanbo Yanjiu* (听见, 看见: 影视媒体的无障碍传播研究) [Hear and See: Audiovisual Media for All]. Hangzhou: Zhejiang Daxue Chubanshe (浙江大学出版社) [Zhejiang University Press].
- Li, D. (李东晓), & Pan, X. (潘祥辉) (2013). Huisheng-huise: Zhongguo wuzhang'ai dianying de fazhan xianzhuang ji zhanwang (绘声绘色: 中国无障碍电影的发展现状及展望) [Vivid and realistic: Development and prospects of Chinese barrier-free movies]. *Zhejiang Xuekan* (浙江学刊) [Zhejiang Journal], 4, 188–198.
- Li, D., & Looms, P. (2016). Television Accessibility in China. In P. Biswas, C. Duarte, P. Langdon, & L. Almeida (Eds.), *A Multimodal End 2 End Approach to Access* (pp. 261–273). Springer.
- Lin, Z., & Yang, L. (2018). Denial of Disability Discrimination by Disabled People in China. *Disability & Society*, 33(5), 804–809. doi: 10.1080/09687599.2018.1446276
- Liu, C. (刘芷晴) (2015). *Ni Shi Wo de Yan. Yunyong Koushu Yingxiang Fuzhu Guoxiao Shizhangsheng Xuexi zhi Yanjiu* (你是我的眼。运用口述影像辅助国小视障生学习之探究) [You are my eyes: Using audio description to assist the learning of elementary school students with visual impairment] (PhD dissertation). National Taiwan Normal University, Taiwan.
- Llisterri, J., Fernández, N., Gudayol, F., Poyatos, J. J., & Martí, J. (1993). "Testing Users' Acceptance of Ciber232, a Text to Speech System Used by Blind Persons". In B. Granström, S. Hunnicut, & K. Spens (Eds.), *Speech and Language Technology for Disabled Persons. Proceedings of an ESCA Workshop* (pp. 203–206).

- Ma, B. (马波) (2016). Qianyi wuzhang'ai dianying (浅议无障碍电影) [On barrier-free movies]. *Dangdai Dianshi* (当代电视) [Contemporary TV], 16, 110–111.
- Ma, W. (马卫) (2019). Wei hei'an shijie zaizao guangying. Mangren wuzhang'ai dianying jiaoben zhuangao chutan (为黑暗世界再造光影——盲人无障碍电影脚本撰稿初探) [Give back the light to a Black world. First exploration in the creation of AD scripts for barrier-free movies]. *Xin Shangwu Zhoukan* (新商务周刊) [New Business Weekly], 255–256.
- Macdonald, S. J., Deacon, L., Nixon, J., Akintola, A., Gillingham, A., Kent, J., Ellis, G., Mathews, D., Ismail, A., Sullivan, S., Dore, S., & Highmore, L. (2018). 'The Invisible Enemy': Disability, Loneliness and Isolation. *Disability & Society*, 33(7), 1138–1159. doi: 10.1080/09687599.2018.1476224
- Mańczyńska, M. (2011). *TTS AD with Audio Subtitling to a Non-fiction Film. A Case Study Based on La Soufriere by Werner Herzog* (Masters dissertation). University of Warsaw.
- Martínez-Sirés, P. (2016). Making the Inaccessible Accessible: A Panorama of Audiovisual Translation, Accessibility and Audio Description in Japan. *Asiadémica*, 7, 24–50.
- Matamala, A., & Rami, N. (2009). Análisis comparativo de la audiodescripción española y alemana de "Good-bye Lenin". *Hermeneus*, 11, 249–266.
- Matamala, A., Orero, P., Rovira-Esteva, S., Casas-Tost, H., Morales-Morante, F., Soler-Vilageliu, O., Agulló, B., Fidyka, A., Segura, D., & Tor-Carroggio, I. (2018). User-centric Approaches in Access Services Evaluation: Profiling the End User. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, & T. Tokunaga (Eds.), *Proceedings of the Eleventh International Conference on Language Resources Evaluation (LREC 2018)* (pp. 1–7).
- Matamala, A.; Machuca, M., & Ríos, A. (2018). More than Words. Voice in Audio Description. Paper presented at Languages and the Media Conference (October 2018), Berlin, Germany. Retrieved from: <https://ddd.uab.cat/record/196126>
- Matamala, A., & Orero, P. (2018). Standardising Accessibility: Transferring Knowledge to Society. *Journal of Audiovisual Translation*, 1(1), 139–154.
- Matamala, A., & Ortiz-Boix, C. (2018). Text-to-speech Voice-over? A Study on User Preferences in the Voicing of Wildlife Documentaries. *Skase. Journal of Translation and Interpretation*, 1, 24–38.

- Mendonza, N., & Matamala, A. (2019). Panorama de la enseñanza de la audiodescripción en España: resultados de un cuestionario. *MonTI. Monografías de Traducción e Interpretación*, 11, 155-185.
- Mitra, S. (2006). The Capability Approach and Disability. *Journal of Disability Policy Studies*, 16(4), 236–247. doi: 10.1177/10442073060160040501
- Morisset, L., & Gonant, F. (2008). La Charte de l’audiodescription. Retrieved from: www.csa.fr [last accessed: 8 July 2019]
- Na, T., & Yu C. (2018). From Multiple Barriers to a Co-prosperity Society: The Development of a Legal System for Disabled People in China. *Disability & Society*, 33(7), 1170–1174. doi: 10.1080/09687599.2018.1471373
- Neves, J. (2005). *Audiovisual Translation: Subtitling for the Deaf and Hard-of-hearing* (PhD dissertation). Roehampton University, UK.
- Neves, J. (2011). *Guia de audiodescrição. Imagens que se ouvem*. Leiria: IPL.
- Neves, J. (2016). Enriched Descriptive Guidelines: A Case of Collaborative Meaning-Making into Museums. *Cultus*, 9(2), 137–154.
- Nikolic, K. (2018) Reception Studies in Audiovisual Translation. Interlingual Subtitling. In E. Di Giovanni, & Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation* (pp. 179–198). Amsterdam: John Benjamins.
- Observatorio estatal de la discapacidad (2018). “Población con discapacidad según zona de residencia”. Retrieved from <https://www.observatoriodeladiscapacidad.info/tag/estadisticas/>
- OFCOM. (2001). Guidelines on the Provision of Television Access Services. Retrieved from: http://audiodescription.co.uk/uploads/general/RNIB._AD_standards.pdf
- Olalla-Soler, C. (2019). Practices and Attitudes toward Replication in Empirical Translation and Interpreting Studies. *Target*, 32(1). doi: 10.1075/target.18159.ola
- Orero, P. (2005). Audio Description: Professional Recognition, Practice and Standards in Spain. *Translation Watch Quarterly*, 1, 7–18.
- Orero, P. (2008). Three Different Receptions of the Same Film. *European Journal of English Studies*, 12(2), 179–193. doi: 10.1080/13825570802151454
- Orero, P. (2012). Audio Description Behaviour: Universals, Regularities and Guidelines. *International Journal of Humanities and Social Science*, 17(2), 195-202.
- 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. *JoSTrans: The Journal of Specialised Translation*, 30, 105–126.
- Orero, P., & Tor-Carroggio, I. (2018). User Requirements when Designing Learning E-content: Interaction for All. In E. Kapros, & M. Koutsombogera (Eds.), *Designing for the User Experience in Learning Systems* (pp. 105–122). Cham: Springer Nature. doi: 10.1007/978-3-319-94794-5_6
- Orrego-Carmona, D. (2015). *The Reception of (Non)professional Subtitling* (PhD dissertation). Universitat Rovira i Virgili, Spain.
- Ortiz-Boix, C. (2012). *Tecnologies per a l'audiodescripció. Estudi sobre l'aplicació de la traducció automàtica i la síntesi de parla a l'audiodescripció en castellà* (Masters dissertation). Autonomous University of Barcelona, Spain.
- Palmer, J. (2014, 17 February). Crippling Injustice. *Aeon*. Retrieved from: <https://aeon.co/essays/what-is-life-like-for-disabled-people-in-china>
- Palomo López, A. (2010). The Benefits of Audio Description for Blind Children. In J. Díaz-Cintas, A. Matamala, & J. Neves, *New Insights into Audiovisual Translation and Media Accessibility* (pp. 213–226). Amsterdam/New York: Rodopi.
- Pappworth, M. H. (1967). *Human Guinea Pigs: Experimentation on Man*. London: Routledge and Kegan Paul.
- Pazos, P. (2012). *Audiosubtitulació: Una possible solució para la accesibilidad a los medios audiovisuales* (Master's dissertation). Autonomous University of Barcelona, Spain.
- Peli, E., Fine, E. M., & Labianca, A. T. (1996). Evaluating Visual Information Provided by Audio Description. *Journal of Visual Impairment & Blindness*, 90(5), 378-385.
- Peng, T. (2018). *Baidu AI Can Clone Your Voice in Seconds*. Retrieved from: <https://medium.com/syncedreview/baidu-ai-can-clone-your-voice-in-seconds-93558a7b984f>
- Peng, X., Song, S., Sullivan, S., Qiu, J., & Wang, W. (2010). Ageing, the Urban-Rural Gap and Disability Trends: 19 Years of Experience in China - 1987 to 2006. *PLoS ONE*, 5(8), e12129. doi: 10.1371/journal.pone.0012129
- People's Republic of China. (n.d.). *Copyright Law of the People's Republic of China*. Retrieved from: <https://www.wipo.int/edocs/lexdocs/laws/en/cn/cn001en.pdf>
- Perego, E. (2019). Audio Description: Evolving Recommendations for Usable, Effective and Enjoyable Practices. In L. Pérez González (Ed.), *Routledge Handbook of Audiovisual Translation* (pp. 114–129). London: Routledge.

- Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The Cognitive Effectiveness of Subtitle Processing. *Media Psychology*, 13(3), 243-272.
- Perotti, O., & Valero Gisbert, M. (2017). Standards for Audio Description. Contrastive Analysis. In A. Bécart, V. Merola, & R. López-Campos Bodineau (Eds.), *Audiovisual Translation. Current Issues and Challenges* (pp. 67-75). Sevilla: Editorial Bienza.
- Petersen, C. J. (2008). China's Ratification of the Convention on the Rights on Persons with Disabilities: The Implications for Hong Kong. *Hong Kong Law Journal*, 38, 611–643.
- Price Waterhouse Coopers (2016). "China media entertainment outlook 2016-2020". Retrieved from <https://www.pwccn.com/en/entertainment-media/em-china-outlook-nov2016.pdf>
- Qiu, Z. (1998). Disability Statistics in the People's Republic of China. *Asia & Pacific Journal on Disability*, 3. Retrieved from <https://www.dinf.ne.jp/doc/english/asia/resource/z00ap/003/z00ap00308.html>
- Rai, S. (2009). *Bollywood for All. The Demand for Audio Described Bollywood Films*. London: Royal National Institute for the Blind.
- Ramos, M. (2013). *El impacto emocional de la audiodescripción* (PhD dissertation). Universidad de Murcia, Spain.
- Ramos, M. (2015). The Emotional Experience of Films: Does Audio Description Make a Difference? *The Translator*, 21(1), 68–94.
- Remael, A. (2012a). Audio Description with Audio Subtitling for Dutch Multilingual Films: Manipulating Textual Cohesion on Different Levels. *Meta: Journal des traducteurs*, 57(2), 385–407.
- Remael, A. (2012b). Media Accessibility. In *Handbook of Translation Studies Volume 3* (pp. 95-101). John Benjamins Publishing Company.
- Remael, A., Reviere, N., & Vercauteren, G. (2015). *Pictures Painted in Words. ADLAB Audio Description Guidelines*. Trieste: Edizione Università di Trieste.
- Reviere, N. (2016). Audio Description Services in Europe: An Update. *JoSTrans: The Journal of Specialised Translation*, 26, 232–247.
- RNIB. (2019). Television. Retrieved from: <http://rnib.org.uk/television>.
- Rodríguez Posadas, G. (2010). Audio Description As a Complex Translation Process: A Protocol. In J. Díaz-Cintas, A. Matamala, & J. Neves (Eds.), *New Insights into Audiovisual Translation and Media Accessibility* (pp. 195–212). Amsterdam/New York: Rodopi.

- Romero-Fresco, P. (2018). Reception Studies in Live and Pre-recorded Subtitles for the Deaf and Hard of Hearing. In E. Di Giovanni, & Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation* (pp. 199–224). Amsterdam: John Benjamins.
- Romero-Fresco, P., & Fryer, L. (2013). Could Audio Described Films Benefit from Audio Introductions? An Audience Response Study. *Journal of Visual Impairment and Blindness*, 107(4), 287–285.
- Rovira-Esteva, S. (2010). *Lengua y escritura chinas: mitos y realidades*. Barcelona: Edicions Bellaterra.
- 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(21), 71–80.
- Sahasrabuddhe, A. (2019). Use of Enacted Audio Description to Improve Film Watching Experience for Visually Impaired Viewers. Paper presented at Media for All 8. Stockholm, 19 June 2019.
- Saldanha, G., & O'Brien, S. (2013). *Research Methodologies in Translation Studies*. S.I.: Routledge.
- Salway, A. (2007). A Corpus-based Analysis of Audio Description. In J. Díaz Cintas, P. Orero, & A. Remael (Eds.), *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 151–174). Amsterdam/New York: Rodopi.
- Sanz-Moreno, R. (2018). El uso de extranjerismos en audiodescripción. La opinión de los usuarios. *Parallèles*, 30(2), 53–69.
- Sanz-Moreno, R. (2019). Competencia cultural del receptor normovente y audiodescripción. Estudio de caso. *HIKMA. Translation Studies Journal*, 18(2), 257–275.
- Schmeidler, E., & Kirchner, C. (2001). Adding Audio Description: Does it Make a Difference? *Journal of Visual Impairment and Blindness*, 95(4), 197–212.
- Snyder, J. (2007). Audio Description: The Visual Made Verbal. *The International Journal of the Arts in Society*, 2(1), 99–104.
- Snyder, J. (2014). *The Visual Made Verbal*. Arlington: American Council of the Blind.
- Stevens, C., Lees, N., Vonwiller, J., & Burnham, D. (2005). On-line Experimental Methods to Evaluate Text-to-speech (TTS) Synthesis: Effects of Voice Gender and Signal Quality on Intelligibility, Naturalness and Preference. *Computer Speech and Language*, 19, 129–146. doi: 10.1016/j.csl.2004.03.003

- Streijl, R. C., Winkler, S., & Hands, D. S. (2016). Mean Opinion Score (MOS) Revisited: Methods and Applications, Limitations and Alternatives. *Multimedia Systems*, 22(2), 213–227.
- Szarkowska, A. (2011). Text-to-speech Audio Description: Towards Wider Availability of AD. *JoSTrans: The Journal of Specialised Translation*, 15, 142–163.
- Szarkowska, A., & Jankowska, A. (2012). Text-to-speech Audio Description of Voice-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.
- Szarkowska, A., & Orero, P. (2014). The Importance of Sound for Audio Description. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio Description. New Perspectives Illustrated* (pp. 121-139). Amsterdam: John Benjamins.
- Szarkowska, A., & Wasylczyk, P. (2018). Five Things You Wanted to Know About Audiovisual Translation Research But Were Afraid to Ask. *Journal of Audiovisual Translation*, 1(1), 8–25.
- The National People's Congress of the People's Republic of China. (n.d.). *Beijing Shi Shishi "Zhonghua Renmin Gongheguo Canjiren Baozhangfa" Banfa* (北京市实施“中华人民共和国残疾人保障法”办法) [Beijing's implementation of the “Law on the Protection of Persons with Disabilities”]. Retrieved from: <http://www.npc.gov.cn/npc/zfjc/cjrbzxfzjc/2012-07/26/content_1731040.htm>
- Thompson, L., Reeves, C., & Masters, K. (1999). In the Balance: Making Financial Information Accessible. *British Journal of Visual Impairment*, 17(2), 65–70.
- Thrane, L. (2013). *Text-to-speech on Digital TV: An Exploratory Study of Spoken Subtitles on DRISyn* (Masters dissertation). University of Copenhagen, Denmark.
- Tor-Carroggio, I. (2018). State-of-the-art of AD Standards in China. Poster presented at Languages and the Media, Berlin, 3-5 October 2018.
- Tor-Carroggio, I., & Casas-Tost, H. (2020). Who is Currently Audio Describing in China? A Study of Chinese Audio Describer Profiles. *MonTI. Monografías de Traducción e Interpretación*, 12, 78–107.
- Tor-Carroggio, I. (2020). The Customer Is Always Right: Study on Chinese Persons with Sight Loss' Opinion on Their Experience with Audio Description. *Disability & Society*. doi: 10.1080/09687599.2020.1713727

- Tor-Carroggio, I., & Vercauteren, G. (2020). When East Meets West: A Comparison of Audio Description Guidelines in China and Europe. To appear in *HIKMA. Translation Studies Journal*.
- Tuominen, T. (2018). Multi-method Research: Reception in Context. In E. Di Giovanni, & Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation* (pp. 69–90). Amsterdam/Philadelphia: John Benjamins.
- United Nations. (2006). *Convention on the Rights of Persons with Disabilities*. Retrieved from: <http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf>
- Vercauteren, G. (2012). A Narratological Approach to Content Selection in Audio Description. Towards a Strategy for the Description of Narratological Time. *MonTI. Monografías de Traducción e Interpretación*, 4, 207–231.
- Vercauteren, G. (2016). A Translational and Narratological Approach to Audio Describing Narrative Characters. *TTR*, 27(2), 71–90.
- Vercauteren, G., & Remael, A. (2015). Spatio-temporal Settings. In A. Maszerowska, A. Matamala, & P. Orero (Eds.), *Audio Description. New Perspectives Illustrated* (pp. 61–80). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- 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* (Masters dissertation). University of Warsaw, Poland.
- Walczak, A. (2017a). *Immersion in Audio Description. The Impact of Style and Vocal Delivery on Users' Experience* (PhD dissertation). Autonomous University of Barcelona, Spain.
- Walczak, A. (2017b). Audio Description on Smartphones: Making Cinema Accessible for Visually Impaired Audiences. *Universal Access in the Information Society*, 17(4), 833–840. doi: 10.1007/s10209-017-0568-2
- Wang, H. (王桦), Zhao, S. (赵晟珣), Zeng, E. (曾尔亢), Ma, C: (马春薇), Wang, J. (王家瑜), & Duan, L. (段凌). (2014). Zhongguo renkou laolinghua shehui fazhan yu yingdui celüe (中国人口老龄化社会发展与应对策略) [The social development and the coping strategies for China's aging population]. *Zhongguo Shehui Yixue Zazhi* (中国社会医学杂志) [Chinese Journal of Social Medicine], 31(2), 75–77.
- Wang, R., Chen, S., Tao, J., & Chu, M. (2006). Mandarin Text-to-speech Synthesis. In C. Lee, H. Li, L. Lee, R. Wang, & Q. Huo (Eds.), *Advances in Chinese Spoken Languages Processing* (pp. 99–124). Singapore: World Scientific.

- Wehrmeyer, J. (2014). Eye-tracking Deaf and Hearing Viewers of Sign Language Interpreted News Broadcasts. *Journal of Eye Movement Research*, 7(1), 1–16.
- Wu, Z. (吴宗艺), & Xie, Z. (谢桢桢) (2015). Zhongguo shizhang koushu yingxiang fuwu de fazhan xianzhuang yu dalu tuiguang (中国视障口述影像服务的发展现状与大陆推广) [The development and the promotion of audio description in Mainland China]. *Xinwen Yanjiu Daokan* (新闻研究导刊) [Journal of News Research], 6(10), 20–22.
- Xiao, W., & Peng, W. (2019). Audiovisual Translation Studies in China. A Long Journey to Academic Acknowledgement. In Li, D., & Han, Z. (Eds.), *Translation Studies in China. The State of the Srt* (pp. 261-280). s.l.: Springer.
- Xinhua News Agency (2019). “Porcentaje de población rural de China reduce en 50% en 70 años”. Retrieved from http://spanish.xinhuanet.com/2019-09/04/c_138363875.htm
- Xu, J. (2018). Zhongguo mangwen tushuguan koushu yingxiangguan bangzhu shizhang renshi “ting” dianying (中国盲文图书馆口述影像馆帮助视障人士“听”电影) [The AD screening room of the China Braille Library helps people “listen” to movies]. *Xinwen Zhishi* (新闻知识) [News], 4, 91.
- Yan, X. (鄢秀), & Luo, K. (罗康特). (2019). Xin shidai de fanyi yanjiu: koushu yingxiang lilun yu shixian de zhongyaoxing (新时代的翻译研究：口述影像理论与实践的重要性) [Translation research in the new era: the importance of AD theory and practice]. *Dongfang Fanyi* [East Journal of Translation] (东方翻译), 4, 8–12.
- Yeung, J. (2007). Audio description in the Chinese world. In J. Díaz-Cintas, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the deaf, audio description, and sign language* (pp. 231–244). Amsterdam: Rodopi. doi: 10.1057/9781137552891_15
- Zhang, C. (2017). ‘Nothing About Us Without Us’: The Emerging Disability Movement and Advocacy in China. *Disability & Society*, 32(7), 1096–1101, doi:10.1080/09687599.2017.1321229
- Zheng, X. (郑玄). (2017). *Estándares de subtitulación de la television china: Un estudio de caso* (Masters dissertation). Autonomous University of Barcelona, Spain.
- Zhou, X. (2002). The Discourse of Disability in Modern China. *Patterns of Prejudice*, 36(1), 104–112. doi: 10.1080/00313220212881139

Annexes

Annex 1: Articles within this dissertation

- 1.1. Article 2: Tor-Carroggio, I. (2020). The customer is always right: Study on Chinese persons with sight loss' opinion on their experience with audio description. *Disability & Society*. doi: 10.1080/09687599.2020.1713727

DISABILITY & SOCIETY
<https://doi.org/10.1080/09687599.2020.1713727>



The customer is always right: Study on Chinese persons with sight loss' opinion on audio description

Irene Tor-Carroggio

Autonomous University of Barcelona, Barcelona, Spain

ABSTRACT

Audio description is in its infancy in China, both in terms of the provision of the service and as an academic object of study. This article provides a useful insight into what Chinese audio description users think of this access service and how they consume it. This information was gathered through two different questionnaires, mostly distributed in the city of Shanghai, although not exclusively. The results show what the profile of the audio description user is, how satisfied they are with the service that is currently available and their expectations for the future. Only after identifying the areas that matter to users the most can both science and industry contribute to guaranteeing their rights in the best possible way.

ARTICLE HISTORY

Received 9 April 2019
 Accepted 7 January 2020

KEYWORDS

Media accessibility; audio description; China; user satisfaction

Points of interest

- Media accessibility, specifically audio description, is in its infancy in Mainland China as an academic object of study. This study contributes to start bridging this research gap.
- Regarding their access to information and culture, the voices of Chinese persons with sight loss are hardly heard. This is the first study that takes into consideration their opinion and experience with audio description.
- The results of this study have had a direct impact on how audio description is delivered in Shanghai.
- The comparison between China and Europe indicates where audio description in China is in relation to countries in which this access service has already come of age.
- The results of this study are key to developing audio description in Mainland China following the United Nation's Convention on the Rights of Persons with Disabilities' slogan: 'nothing about us without us'.

CONTACT Irene Tor Carroggio irene.tor@uab.cat Department of Translation and Interpretation and East Asian Studies of the Autonomous University of Barcelona
 © 2020 Informa UK Limited, trading as Taylor & Francis Group

1. Introduction

Confucianism, Buddhism and Taoism have played a key role in shaping the concept of disability in China. In early China physical abnormality was not inexorably identified with negativity, as popular deities often had uncanny body shapes (Zhou 2002, 105). Also, the *Zhuangzi* contains some passages that suggest that physical disability was not perceived as negative (Zhou 2002; Lewis 2014; Lambert 2016). Yet, this changed and throughout history Chinese persons with disabilities have been seen as not worthy of attention, as the result of karma (Campbell and Uren 2011) or their parents' defects (Palmer 2014), as a deficient fusion between man and nature (Avery 2016), as criminals, as outsiders and even as racial degeneration and one of the key causes for the nation's backwardness (Zhou 2002), just to name a few. It is, therefore, not surprising that persons with disabilities try to run away from traditional stigmatization. After interviewing persons with disabilities in China, Lin and Yang (2018) concluded that their interviewees shared a common feature: they denied being victims of any discrimination and, thus, did not pursue social inclusion. According to these researchers, this phenomenon is caused by a host of institutional, environmental, social, family-related and psychological factors, most of which have been inherited from previous generations' prejudices.

Like in any other language, the choice of words in Chinese reflects social attitudes towards certain issues, among which disability can be found. In China, the terms referring to it have negative connotations, although Aguado Díaz (1995, 277) observed that some social practices carried out in the past, like foot binding, evidence the cultural relativity of physical disability and its context-sensitiveness. Currently, the most widespread word to refer to persons with disabilities is *canjiren* (残疾人), which literally means 'a person who is deficient and ill'. Other words that clearly exemplify how medicalized the conception of disability is in China are, for example, *canfei* (残废, 'crippled') and *shazi* (傻子, 'simpleton'). Yet, new voices have emerged that advocate for new terms to refer to those functionally diverse, such as *canzhang renshi* (残障人士), 'a deficient and handicapped person' (Hallett 2005; Palmer 2014). Unfortunately, although it sounds more formal, it still frames disability within the medical model of disability, which believes that disability stems from biological problems that must be addressed and that only concern individuals. This model has already been condemned by, for instance, the International Telecommunication Union (2017), which endorses an approach that vindicates the idea that the problem emanates from social behavior, instead of the person with an impairment (Orero and Tor-Carroggio 2018).

The establishment of the China Disabled Persons' Federation by the Chinese central government in 1988 was the first step towards the

development of policies and legislation on disability (Zhang 2017). Also, since China's ratification of the United Nations' Convention on the Rights of Persons with Disabilities in 2007, the country has made visible efforts to improve the life of persons with disabilities, such as the amendment of the Law on the Protection of Persons with Disabilities, originally passed in 1980. This is no surprise taking into account that "the occurrence of important domestic and international events often becomes a critical policy window for the improvement of the welfare of disabled people" (Tang and Cao 2018, 1173). Thus, the almost 83 million persons with disabilities in China are also benefiting from the country becoming more aware of their needs. One of the rights the United Nations' Convention on the Rights of Persons with Disabilities recognizes—hence the need of China to give an answer to it—is the access to culture and information. In the case of persons with sight loss, who make up the group this paper revolves around, this right can be granted through sensory accessibility services such as audio description. Audio description, an exponentially growing Audio-visual Translation mode, is a means of translating visual elements into aural description primarily for helping persons with sight loss to enable them to access live and pre-recorded programs and performances. In other words, it is the visual made verbal (Snyder 2014).

In fact, persons with no sensory loss receive information from two sensory modalities: sight and hearing. Audio description interacting with the original sound track 'should stimulate a number of other senses in the user' (Fryer 2016, 47) and therefore grant users a similar experience to that of sighted viewers. Some users have even argued that their experience is enhanced because the lack of visual reminders, such as a TV screen, eradicates the need for suspension of disbelief (Fryer 2016, 47). This accessibility service is, along with others such as subtitles for the deaf and hard of hearing and sign language interpreting, already legally required in countries such as Spain, the UK, and even Japan—although to a different extent (Martínez-Sirés 2016, 34).

In contrast, audio description is still at its infancy in China, both as a professional service and as a research topic (Tor-Carroggio and Casas-Tost, forthcoming). Martínez-Sirés (2016, 27) had also highlighted the lack of scholarship in media accessibility—an umbrella label for a number of sensory access services—for non-Western languages. Known as *koushu yingxiang* (口述影像, 'image description') and *wuzhang'ai dianying* (无障碍电影, 'barrier-free movies'), depending on the province and on the level of accuracy aimed at, some researchers have tried to contextualize audio description from a historical point of view (Li 2013) and even by reporting its state of the art and profiling the persons in charge of delivering it in China (Tor-Carroggio and Casas-Tost, forthcoming). Nonetheless, one of the main angles from which audio description needs to be studied remains unveiled: user needs and satisfaction. In fact, Matamala and Orero (2018)

consider this to be of the utmost importance and underline the need of making the end user's voice heard, and that research supports the statements they make.

This paper aims at bridging this existing research gap by presenting the results of two different surveys conducted in Mainland China in 2018. Both questionnaires departed from the following three research questions:

1. What are the habits of audio description users when it comes to this accessibility service?
2. What are the needs of audio description users when it comes to this accessibility service?
3. How satisfied are users towards the audio description currently delivered?

This article is divided into four sections. The first one revises previous studies that have focused on Chinese persons with sight loss and Chinese media accessibility. The second one explains the methodology followed in the two studies presented in this paper. The third section analyzes the results of the two surveys conducted with audio description users. Finally, the discussion and conclusions are put forward.

2. Research background

Although the events related to it are reported in the press from time to time, media accessibility is still an under-researched field in China (Gambier and Jin 2018). This is also proved by doing a quick search combining 'Chinese' and 'media accessibility' in the Translation Studies database BITRA, which is the field, more specifically Audiovisual Translation, in which media accessibility is encompassed. Currently, such a search only yields four results. Something similar happens when consulting the Chinese academic database CNKI. Consequently, it is not surprising that only a few scholars have investigated audio description users to find out what their needs are and how audio description should be delivered to suit them best. Chao (2002) and Liu (2015) carried out some research in Taiwan in this regard. The former was the first academic researcher to consider audio description as a research topic in Greater China, while the latter studied the need of using audio description to assist the learning of elementary school students with visual impairment. This is also the case of Leung (2018), who examined the media use, behaviour and motivations, as well as the reception and preferences, of visually impaired audiences when consuming audio description in Hong Kong. Similarly, Li (2013), in a rather comprehensive study on media accessibility in Mainland China, included a section presenting the results of a survey administered to persons with sight loss ($N = 729$) in Zhejiang province from which their media use, behavior, and satisfaction were elicited. Her research

confirmed that the most popular means of communication amongst them were television (86.6%), radio (75.3%), mobile phones (68.1%), the Internet (45.3%), DVD/VCD (24.8%), movies (21.2%), and newspapers (20.5%). The respondents claimed to use the media primarily to get informed (64.0%) and to entertain themselves (41.3%), and they were 'quite satisfied' (51.6%) with how it met their needs. As for their habits regarding the TV, 50.0% of the informants watched it more than one hour every day and the most popular type of program was the news. Li (2013) connected the latter finding with the fact that news programs allow persons with sight loss to be less dependent on the images being broadcast. The fact that audio description is not incorporated in TV programs did not prevent Li's (2013) respondents from affirming they were satisfied with TV content (Li 2013, 59–60). As for the lack of audio description on TV, Li and Looms (2016) investigated the current incompatibility between Chinese TV sets and audio description:

One of the reasons for not including the capability to offer closed captions or audio description in China's digital television standard, China Multimedia Mobile Broadcasting (CMMB) 10 years ago was a concern about the affordability of set-top boxes (STBs) [sic]. Industry sources suggest that the retail price of STBs had to be kept below RMB 200 (currently USD 30). As the inclusion of closed captioning and audio description would have required patent licenses, this would have driven costs above the target retail price (Li and Looms 2016, 267–268).

As for movies, which are the only field in which audio description is applied in China, Li's (2013, 41–47) results show that they are a less popular choice amongst blind and partially-sighted persons. More than half of the respondents (54.0%) never watched movies mainly because of the inconvenience of getting to the venue.

The scarcity of academic bibliography might derive from society's general lack of interest in this group of persons, as the results from Wu and Xie (2015) study revealed. These researchers conducted an online survey in China and found out that, out of the 160 valid responses gathered, 79.4% of the respondents admitted not caring about persons with sight loss. This is probably why 88.8% of them did not know about the existence of audio description (Wu and Xie 2015). The lack of literature can also be attributed to persons with sight loss not having attached enough importance to or not having vindicated enough their right to culture and information because of them having to deal with other difficulties. In fact, Li's (2013) survey proved that the main challenges respondents had to face were poverty (24.7%) and loneliness (24.7%). She even recounted in her book how users questioned her research for not focusing on what they considered more critical problems, such as employment. Also, the scant research in this field has not been compensated from abroad, as, as far as we know, only Fei (2011) wrote a masters dissertation on audio description and the Chinese community in

Canada, which aimed at examining the differences and similarities of expectations and evaluations of integrated audio description between the Chinese and Canadian communities. Yet, his results need to be contextualized, since the stimuli he used was audio description in English with both groups.

It is in this context of lack of studies on audio description in China that this researched is framed in, with the aim of making a step forward from the users' perspective.

3. Methodology

This research presents the results of two different questionnaires that catered for the same target respondents—Chinese persons with sight loss that had experience with audio description.

The first questionnaire was inspired on an existing one created by the European project ADLAB PRO (<http://www.adlabproject.eu/>), which gauged user satisfaction and opinions on audio description. It was 'inspired' and not 'based' on that survey because numerous changes had to be made to fit the Chinese case. For example, 'opera' was replaced for 'Chinese opera' in those questions in which some fields were listed so that users chose where they wanted to enjoy audio description. The questionnaire had also to be shortened because the circumstances in which the questionnaire was expected to be delivered would not allow us to enter into much detail. The questionnaire was then orally piloted in a blind massage parlor on 18 April 2018 in Shanghai. Five blind masseurs were recruited and only one had previous experience with audio description. Yet, all their comments provided useful insight into the design of questionnaires for this respondent profile. The pilot resulted in many changes, such as adding a clearer explanation of what audio description was in the information sheet and making all the information and questions more easy-to-read. This was attributed to the users' low educational level and lack of knowledge about this accessibility service. It was also found that these users did not know the Chinese term for audio description and an explanatory remark was made to show that the access service delivered by the so-called 'barrier-free movies' was also known as *koushu yingxiang*. Also, it was confirmed that users had difficulty in understanding 1–10 scales and were consequently substituted by categorical ones. All these changes arising from the pilot were considered to be a normal practice, as, for instance, Chmiel and Mazur (2012) indicated they had to change their testing materials up to sixteen times before formally administering them. Before conducting our survey, and since applying for ethical approval is becoming commonplace in the execution of Audio-visual Translation research (Orero et al. 2018), our questionnaire was approved by our university's Ethics Committee. Ethical considerations in human research have been a major concern since the critical articles written by Pappworth

(1967) on medical research. The final questionnaire was administered at the Cathay Cinema in Shanghai on 26 April 2018. This cinema is one of the 16 cinemas that offer a live audio description session for free thanks to the financial help of the local government (Tor-Carroggio and Casas-Tost, forthcoming). With the help of some volunteers, who had been briefed on what audio description is and also on the aim of the research, cinema-goers with sight loss were approached while they were lining up to get into the cinema for their monthly live audio description session. This was thought to be the best way to find as many users as possible, since one of the requirements to answer the questionnaire was to be or to have been an audio description user. An online version of the questionnaire was also made available in the website Web Survey Creator. It was sent to the China Braille Library in Beijing so that they helped us recruit informants from a different city. Therefore, the survey was administered orally and through the Internet.

The second questionnaire was both designed and administered by the association Sound of Light in Shanghai. Sound of Light is an association led by Han Ying which was set up in Shanghai in 2016. It receives direct material and financial assistance from the Shanghai Association of Persons with Disabilities. They produce about 50 audio described movies every year, which are recorded in their headquarters and distributed through digital means. The movies are audio described thanks to a network of volunteers that both write the audio description script and voice it. These movies can be accessed at specific projecting sites which are all located in Shanghai (Tor-Carroggio and Casas-Tost, forthcoming). The aim of the questionnaire was to gather data about the experience, if any, of persons with sight loss regarding the so-called 'barrier-free movies', as well as their opinion on them. In this case, the survey was only conducted in Shanghai. It was administered in person and orally by volunteers and by the association's staff before or after the audio description sessions. The surveys were collected by Sound of Light employees and stored in the association's headquarters. Data collection started in April 2018 and was extended until January 2019.

The reasons for presenting the results of both surveys in the same paper are the following. First, the first questionnaire's sample was smaller than the second one, so it was thought that the first findings could be contrasted and complemented with the data obtained from the second survey. Yet, the first questionnaire could not be replaced by the second one because the questions asked were not the same and both provided significant information about user experience and satisfaction towards audio description. Second, despite the fact that the second questionnaire was not designed by myself, I was offered the data to analyze it in order to ultimately present the results to the Shanghai Association of Persons with Disabilities. Finally, as it has already been underlined, this is the first time this kind of data from

Mainland China is being presented to academia and the aim of the paper was to compile what is currently available and offer as complete of a picture as possible of what has been done so far in this regard.

For the statistical analysis IBM SPSS (version 22) was used.

4. Results of survey 1

The first questionnaire that was administered was framed within PhD research aimed at verifying if Chinese persons with sight loss would accept the use of artificial voices in the audio description of films. Before doing so, audio description users had to be studied and profiled to better design the next stages of the PhD research.

4.1. Informants' demographic profile

Fifty-two informants answered this first questionnaire. They were from both Shanghai (42) and Beijing (10). The average age was 63.8 years old and the median was 65 years old. The youngest informant was 28, while the eldest was 88. It is worth highlighting that 17 persons decided not to answer this question. In the light of the average age observed, it was not surprising that 88.5% of our respondents were retired. Almost 6% of our respondents were unemployed, 1.9% were freelancers and 3.8% of the persons who answered our questionnaire stated that they had a different employment situation than those listed.

Figure 1 clearly shows that only a minority (3.8%) had attended university, with secondary education being the educational level which the majority of informants had completed.

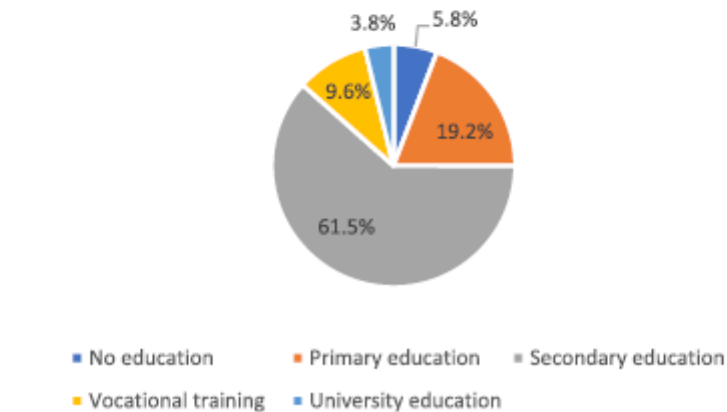


Figure 1. Informants' educational background.

As regards their mother tongue, more than half of the respondents (57.7%) claimed to speak a Chinese dialect, although Standard Chinese (42.3%) is the language in which movies are mainly audio described in China. The average level of Standard Chinese was found to be 0.83 out of 2. A reason that could explain our finding could be the well-known Chinese modesty and cautiousness. Yet, Rovira-Esteva (2010, 257) also noted that, according to official data from 2004, the older Chinese persons are, the less they can use Standard Chinese to communicate. The same happens with their academic background: the more educated they are, the better they can communicate in the country's *lingua franca*.

With respect to their visual condition, 67.3% of the informants claimed that their sight loss was congenial, whilst 23.1% said it was acquired.

4.2. Informants' experience and opinion on audio description

All the informants were consumers of audio described films, and more than 80% did so in cinemas, but none of them selected any of the other areas in which audio description can be provided, i.e., TV, museums, among others. Yet, they would be interested in having the service available in different fields, as Figure 2 points out.

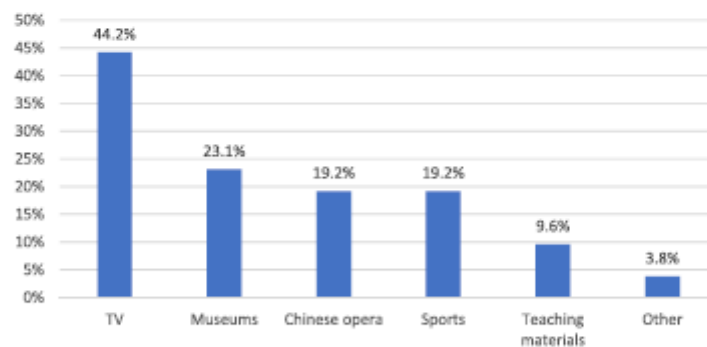


Figure 2. Fields in which respondents would like to have AD.

TV revealed itself as the most sought-after mode for audio description, although the distribution obtained was not statistically significant ($\chi^2 = 0.692$, $p = 0.405$). This meant that a distribution in which 'yes' and 'no' had the same weight could not be discarded.

As for the frequency of use of the service, three quarters of the informants claimed to consume audio description in films frequently or very frequently. 15.4% did so sometimes and only 9.6% of the informants declared that they rarely watch audio described films.

One of the main points of the questionnaire was to determine the degree of user satisfaction with audio description. It was observed that our informants were more satisfied with the quality of audio description (mean = 3.33, in a 0–4 scale) rather than with its quantity (mean = 3.15, in a 0–4 scale). In order to determine if satisfaction in terms of quantity was statistically different to that in terms of quality, a Wilcoxon signed ranks test was carried out. In spite of the means showing some slight differences, it is not possible to state that users are more satisfied with audio description quality than with the quantity offered ($Z = -1.415$, $p = 0.157$, with 33 ties).

It was also deemed interesting to see if there was any correlation between the frequency of audio description consumption and user satisfaction, both in terms of quality and quantity. A statistically significant lineal relation was found—although weak—between audio description consumption frequency and user satisfaction with quantity ($r_s = 0.437$, $p < 0.05$).

The last question sought the opinion of users about which types of movies they considered more suitable to carry out some experiments with artificial voices. Although this experiment is not the focus of this paper, the results of this question are also added because they can be understood as the genres users liked the most. Respondents could select more than one option. The top three choices were historical (36.5%), war (25.0%), and inspirational (25.0%) movies (Figure 3).

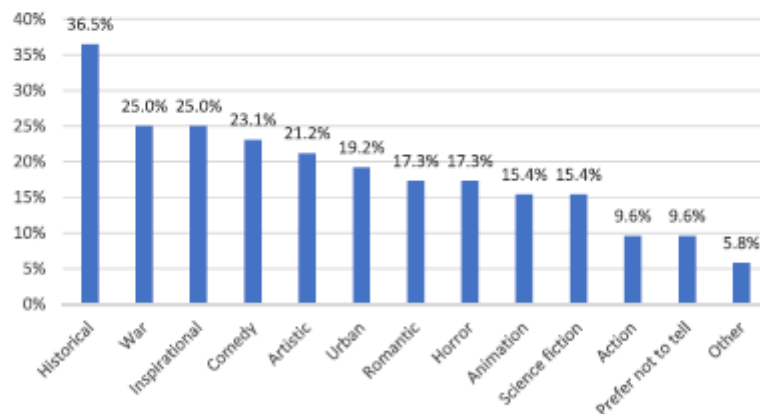


Figure 3. Informants' preferred movie genres.

5. Results of survey 2

The first section presents their demographic profile, whereas the second one focuses on their experience and opinion on audio described movies.

5.1. Informants' demographic profile

One hundred and fifty-six persons answered this questionnaire. The proportion of women (56.4%) was slightly higher than that of men (41.7%). Just about half of the informants (50.6%) were 60 years old or above. Less than 1% were 18–25 years old, those aged 18–35 only represented 4.5% of the respondents, the persons aged 36–45 represented 12.8% of the respondents and almost a third were 46–59 years old (30.1%). Three persons did not answer this question.

As for the informants' academic background, the vast majority did not attend university (92.9%), as Figure 4 shows. This is in line with the age of most informants, who probably did not have a chance to access even basic education in their youth because of their physical impairment or the socio-political situation in the country at that time.

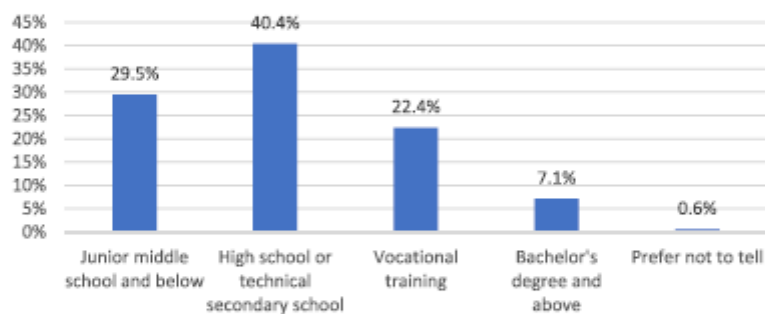


Figure 4. Informants' education background.

Regarding their employment situation, it can be stated that there were no students among the informants. More than half of the respondents were retired (66.7%). The rest were either employed or depended on their families (28.2%) or had other situations (4.5%).

5.2. Informants' experience and opinion on audio description

Regarding whether informants were aware of the existence of audio described movies, 90.4% of the respondents claimed to know about it, whereas 12 persons (7.7%) admitted they did not know and three persons refused to answer this question. This result makes sense given the fact that those surveyed were recruited during activities related to the so-called 'barrier-free movies'.

When asked about how they came to know about this service, more than 80% revealed that they had learned about it through user associations. The rest of the informants mentioned 'assistive staff' (23%), 'family and friends' (16%), and 'the media and the Internet' (15%). Twenty persons did not answer this question.

Regarding whether respondents had the need of audio described movies, an overwhelming majority thought so (94.2%) and only three persons preferred not to answer. Those who answered 'no' in this question were asked to stop filling in the survey.

The remaining 148 informants were then asked what their purpose was when watching audio described movies. They were allowed to select more than one option. The choices given were 'for entertainment' (67.3%), 'to socialize' (37.8%), and 'for study purposes' (49.4%).

Informants were also asked if they had ever watched movies with audio description. Almost 84% of the respondents had watched them (83.3%). These persons were mostly satisfied or very satisfied (76.3%) and only one person admitted not being satisfied with them.

The questionnaire was also interested in knowing how respondents got to the place where barrier-free movies are screened. More than half of the 128 respondents that answered claimed to use public transport, such as the bus or the metro (64.1%). Less than a quarter of respondents (17.3%) stated they went on foot and one informant chose the option of 'others'.

When asked whether there were parts of the audio described movies that were not understood while watching them, more than half of the respondents said that there were almost no such parts (55.1%). Some respondents admitted that they sometimes had to face this situation (26.9%) and only one person admitted that this situation happened quite frequently. Twenty-seven informants preferred not to answer.

To the question about how users wanted the information regarding barrier-free movies to be delivered to them. The option of reaching out to them through Wechat or QQ, both instant chat tools, was the most popular one (30.8%) among the 125 informants that answered the question, closely followed by a Wechat public account (26.9%). Door-to-door notification was only chosen by 16.7% of the informants and nine persons selected the option of 'others'.

Informants were asked what kind of movies they liked the most. They were able to choose three of the 12 types presented. The top five were comedies (47.4%), historic films (37.8%), inspirational films (33.3%), romantic films (27.6%), and war films (25.6%), as Figure 5 shows.

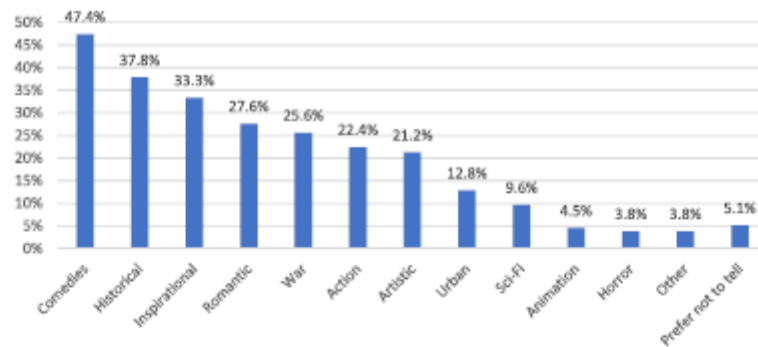


Figure 5. Informants' preferred movies.

Participants were also asked what type of movies they preferred: classic or new. The option of 'both' was also available. Almost 40% of the respondents (39.7%) said that they liked classic movies the most, whereas only 14.1% stated they enjoyed newer movies more. Still, 37.2% of respondents affirmed they liked both types. Fourteen persons refused to answer.

Regarding this last question, an association between the age and the type of movie chosen was found to be statistically significant ($\chi^2 = 20.39$, $p = 0.009$). This means it can be affirmed that, in general, the older a person is the more they prefer to watch old films. In our case, it was observed that persons who were 60 or older preferred old movies.

Finally, a qualitative question was added to allow informants to make any remarks they thought were appropriate. First, some respondents thought that audio described movies needed to be screened more often and better publicized so that more users could benefit from them. Second, it was considered necessary to turn down the volume of the original soundtrack because it sometimes did not allow users to hear the audio description clearly. Third, a few informants wanted the types of movies screened to be more varied and to include more foreign films. Fourth, some persons also asked for more staff to help users when entering and leaving the facilities. Finally, many users hoped that in the future persons with and without sight loss could watch movies together at the cinema. They thought this could be achieved by persons with sight loss using headphones to listen to the audio description.

6. Discussion

These two surveys have helped us draw a picture of the users' profile of audio description in China. Although in both cases the sample was quite aged, being a limitation of this study, the users' similar profile allowed us to draw justifiable comparisons among the two questionnaires. Also, despite

the sample undoubtedly not being representative of the whole China, it must be stated that vision problems are mostly age-related (Fryer 2016, 43), so the samples, even though they can be criticized, can also shed some light on the majoritarian group of audio description consumers. It could be deduced from the average age in both questionnaires that audio description is mostly consumed by elderly persons, but the truth is that audio description sessions for children are also carried out and recorded material is produced for them as well. Fryer (2016, 43) also claims that women are more likely than men to become blind simply because they live longer. This is supported by the slightly higher proportion of women in the sample recruited by Sound of Light. The first questionnaire did not ask about informants' sex because it was not deemed necessary to assess user experience. The demographic similarities between the two samples go beyond age, since in both cases more than 90% of the informants did not attend university and the clear majority were retired.

All the informants were consumers of audio described movies and none mentioned any kind of experience with other types of audio description. Leung's (2018) results also point out that audio description is mostly consumed in films in Hong Kong, although it has a wider scope in the region and, therefore, can also be found in theatres and art exhibitions, for example. All our informants watched audio described movies frequently, which can be explained by the fact that their age allows them to have quite a lot of free time. In contrast, Leung's (2018) informants did so mostly sometimes (38.6%) or rarely (27.3%). Thanks to the Sound of Light questionnaire we found out that users perceived audio description mainly as an entertainment tool. This is in line with Leung's (2018) informants, who also claimed to watch films to entertain themselves. Yet, the respondents of the Sound of Light questionnaire also revealed having other objectives when watching audio described movies, such as socializing and learning, which is line with studies carried out in other countries, such as that of Fryer (2016, 49), who had already pointed out that users valued audio description as a means of social inclusion. In fact, Li (2013) had already found out that one of the problems persons with sight loss had to face the most was loneliness, along with others such as poverty. Macdonald et al. (2018) framed persons with disabilities' loneliness in the social model of disability and argued that so many of them feeling that way was due to the presence of certain disabling barriers, not because of their impairment.

As for user satisfaction with the audio description available so far, it seems that they think quite highly of it, which might seem strange taking into account that mainly volunteers with no professional training deliver it (Tor-Carroggio and Casas-Tost, *forthcoming*). This can be understood in different ways. Since persons with disabilities have traditionally been ignored and not

taken care of in China, it is possible that any initiative that aims at helping or including them is highly praised, no matter the level of professionalism with which it is carried out. Yet, Branje and Fels (2012) found out that amateur description could maintain an acceptable level of quality so the possibility that the audio description prepared in China being of good quality—despite its conditions—is also more than feasible.

Last but not least, users preferring historical films in general was probably related to their age, since the top ten highest-grossing 2018 films in the Chinese market were consulted and none belonged to that genre. Similarly, Leung (2018) found out that historical movies were the second preferred genre in her study, in which 44 persons with sight loss, mostly aged above 40, participated. Therefore, in case this information is used for the preparation of future experiments or even for the real delivery of the service, it must be taken into account that the movie genre probably cannot be divorced from the age of participants.

6.1. The European case

Since the first questionnaire presented was originally used in the ADLAB PRO project, it is possible to compare our results with some of those obtained in the European case. In spite of the respondents coming mostly from 23 European countries, some non-European respondents coming from, for example, Brazil and Mexico were recruited as well (ADLAB PRO 2017). Although the European results do not represent the whole continent, the picture is still quite valuable because the most represented countries in the survey (mainly the UK, Spain and Italy) are those in which audio description has already come of age.

According to ADLAB PRO (2017), the original questionnaire was drawn up in all the project languages, i.e., English, Italian, Polish, Spanish, Slovene and Dutch. One hundred audio description users responded the questionnaire but, unlike the respondents from our sample, not all of them were persons with sight loss (5%). This detail also illustrates the fact that in Europe it is already common to believe that, although mainly targeted at persons with sight loss, audio description can also benefit sighted viewers. This is even endorsed by the International Organization for Standardization's Technical Specification 20071-21 'Information Technology — User interface component accessibility — Part 21: Guidance on audio description' (ISO (International Organization for Standardization) 2014).

The sample recruited was more representative than the ones presented in this paper in terms of age, since they managed to find a more balanced number of respondents for each age bracket. Also, it is worth mentioning that 51% of the respondents had attended university. This stands in stark

contrast with the educational background of our Chinese informants, almost none of whom had attended university. Had the Chinese samples included younger respondents, this could have been different, since Tang and Cao (2018, 1172) state that after cases that hit the headlines such as that of Dong Lina, education institutions are committed to provide special arrangements for disabled people when they participate in any national examinations according to the newest revision of the People with Disabilities Education Ordinance.

Unlike our informants, the ADLAB PRO respondents admitted having access to more kinds of audio description, although most use it to watch films or TV, or when visiting a museum. Audio description is not used that often in the case of theatre, opera and other live events. Some Slovenian respondents complained about the inconvenience of the time programmes with audio description are broadcast or even the lack of it in theatres or opera. However, it can be observed that China is lagging behind because audio description is still not even offered on TV, which is what users demand the most and is probably the most popular medium for audio description in many countries such as Poland (Jankowska and Walczak 2019).

Regarding end-user satisfaction, the European results showed that most users were dissatisfied (46%) or very dissatisfied (22%) with the quantity of audio description provided in their country, although they were quite satisfied with the quality of the products that were offered (14% were very satisfied and 58% were satisfied). We also observed users being more satisfied in terms of quality, but they also claimed to be quite satisfied with the quantity provided, probably because they are still not aware of all the applications audio description can have.

The original ADLAB PRO questionnaire inquired about the aspects related to audio description that users appreciated and disliked the most in regards textual and technical aspects, as well as audio description language and style. Yet, due to the conditions under which the first questionnaire circulated, this kind of questions were left out. Focus groups or one-to-one interviews could be arranged in the future to go into these areas more in depth.

7. Conclusions

Chinese persons with disabilities rarely raise their voices to express themselves or to get engaged in rights advocacy (Zhang 2017). Yet, there are already some accessibility services mostly targeted at persons with hearing and sight loss with which access to information and culture, a right recognized in the United Nations' Convention on the Rights of Persons with Disabilities, can be granted. Audio description is one of these services and in order to enhance its provision and academic research, which are rather

scarce in China so far, users need to be involved. Not much is known about Chinese persons with sight loss, but these are key to developing audio description and making sure the service meets their needs.

This article is the first of its kind, since it provides a deep insight into what Chinese audio description users think of the service and how they consume it. Only after identifying the areas that matter to users the most can science and industry contribute to guaranteeing their rights in the best possible way. Some conclusions can be drawn from the results of the two questionnaires carried out. First, old persons with little or no education and who are already retired seem to be the users that most attend audio description sessions in the cities where the sample was taken, Shanghai and Beijing. Second, they only have access to audio description in films, which they consume quite frequently mainly for entertainment purposes and thanks to the information provided by user associations. Third, they are all satisfied with the service provided so far, despite it not being delivered by persons who have been professionally trained in this field. Fourth, their favorite movie genres seem to be historical films and comedies.

Finally, it is worth highlighting that the results of the Sound of Light questionnaire had immediate consequences that had an impact on the audio description policy of the association. For example, in light of the fact that most of their users are elderly persons with little education, they came up to the conclusion that the terminology in the scripts needed to be simplified. This could actually be a future line of research to explore, since, to the best of our knowledge, easy reading and audio description have never been combined before, although there are some EU-funded projects that have already started looking into that direction, such as the EASIT project (<http://pagines.uab.ca/easit/en>). Also, they decided to take into account the movie genres preferred when selecting the movie to be audio described.

Our study is not exempt of limitations. First, the first questionnaire was delivered in rather difficult circumstances, in which other cinema-goers would be expressing their opinion while one was being surveyed. This may have prevented the informant from expressing their true views by exposing them to peer pressure. Second, since Chinese volunteers helped with the administration of the first questionnaire, and although they were briefed on our research, one cannot be sure that all the questions raised by the informants were answered appropriately. Third, in both cases, especially the first one, the sample was rather small and was not obtained randomly but by recruiting persons attending the screenings of audio described films. Also, the respondents in both cases were mostly elderly persons coming mainly from one Chinese city. Therefore, we cannot draw conclusions for the whole of China and future studies should include bigger and aleatory samples, integrated by persons coming from a wider range of places in China.

Although Chinese persons with sight loss have little experience with audio description, it is worth carrying out research to elicit their opinions on it, as Chmiel and Mazur (2012) underlined when mentioning countries in which audio description still had a lengthy path ahead of it. Also, working within the United Nations' Convention on the Rights of Persons with Disabilities paradigm means to consult end users, following the 'nothing about us without us' approach. That is the reason why, in case audio description wants to be enhanced in Mainland China, the academic research concerning it must frame itself to a user-centric framework.

It is hoped that this paper will raise awareness on how important it is to involve persons with sight loss in the provision or research of anything that concerns them directly. Li and Looms (2016) assured removing the barriers persons with sight and hearing loss face clearly has a high priority in China, so it is expected that studies like this will set the course for future research because the need for it is overriding.

Acknowledgements

The author is member of the TransMedia Catalonia research group (2017SGR113) and is currently enrolled in the PhD Programme in Translation and Intercultural Studies at the Autonomous University of Barcelona. The author would like to thank all the Chinese persons with sight loss taking part in her study.


Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This article has been funded by the EasyTV project (GA761999), the RAD project (PGC2018-096566-B-I00) and the Confucius Institute.

ORCID

Irene Tor-Carroggio  <http://orcid.org/0000-0003-2924-014X>

References

- ADLAB PRO. 2017. "Audio Description Professional: Profile Definition." Accessed 2 September 2019. <https://www.adlabpro.eu/wp-content/uploads/2018/04/IO2-REPORT-Final.pdf>
- Aguado Díaz, Antonio León. 1995. *Historia de las Deficiencias*, 40–41. Madrid: Escuela Libre Editorial, Fundación ONCE.

- Avery, Kevin. 2016. "Disability and the Three Traditional Chinese Belief Systems". Accessed 7 November 2018. <https://www.chinasource.org/resource-library/articles/disability-and-the-three-traditional-chinese-belief-systems>
- Branje, Carmen, and Deborah Fels. 2012. "LiveDescribe: Can Amateur Describers Create High-Quality Audio Description?" *Journal of Visual Impairment & Blindness* 106 (3): 154–165. doi:10.1177/0145482X1210600304.
- Campbell, Anne, and Marie Uren. 2011. "The Invisibles" ... Disability in China in the 21st Century." *International Journal of Special Education* 26 (1): 12–24.
- Chao, Ya-li (赵雅丽). 2002. *Yuyan Shijie Zhong de Liudong Guangying. Koushu Yingxiang de Lilun Jiangou* (语言世界中的流动光影. 口述影像的理论建构) [The moving lights in the world of language. Theory of audio description]. Taipei: Wunan Publishing Co.
- Chmiel, Agnieszka, and Iwona Mazur. 2012. "AD Reception Research: Some Methodological Considerations." In *Emerging Topics in Translation: Audio Description*, edited by Elisa Perego, 57–74. Trieste: EJT.
- Fei, Xue. 2011. "Audio Description in the Chinese Community." Masters diss., Capital University of Economics and Business, Canada.
- Fryer, Louise. 2016. *An Introduction to Audio Description*. London: Routledge.
- Gambier, Yves, and Haina Jin. 2018. "Audiovisual Translation in China: A Dialogue between Yves Gambier and Haina Jin." *Journal of Audiovisual Translation* 1 (1): 26–39.
- Hallett, Stephen. 2005. "One Eye on China: Challenging Disability". Accessed 2 November 2018. <http://www.bbc.co.uk/ouch/features/one-eye-on-china-challenging.shtml>
- ISO (International Organization for Standardization). 2014. "TS 20071-21 'Information Technology — User Interface Component Accessibility—Part 21: Guidance on Audio Description'".
- International Telecommunication Union. 2017. "Question 7/1: Access to Telecommunication/ICT Services by Persons with Disabilities and with Specific Needs." Accessed 12 October 2018. https://www.itu.int/dms_pub/itu-d/opb/stg/D-STG-SG01.07.4-2017-PDF-E.pdf
- Jankowska, Anna, and Agnieszka Walczak. 2019. "Audio Description for Films in Poland: History, Present State and Future Prospects." *Jostrans: The Journal of Specialised Translation* 32: 236–261.
- Lambert, Andrew. 2016. "The Usefulness of the Useless': Rethinking Disability through Classical Daoist Thought." In *Disability and World Religions: An Introduction*, edited by Darla Y. Schumm and Michael Stoltzfus. Waco: Baylor University Press.
- Leung, Hoi Ching Dawning. 2018. "Audio Description of Audiovisual Programmes for the Visually Impaired in Hong Kong." PhD diss., University College London.
- Lewis, Grison. 2014. "Use of the Useless': Assessing Depictions of Disability in the Zhuangzi." Bachelor's diss., Brigham Young University. Accessed 20 November 2018. https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1017&context=libraryrg_studentpub
- Li, Dongxiao (李东晓). 2013. *Tingjian, Kanjian: Yingshi Meiti de Wuzhang'ai Chuanbo Yanjiu* (听见, 看见: 影视媒体的无障碍传播研究) [Hear and See: Audiovisual Media for All]. Hangzhou: Zhejiang Daxue Chubanshe (浙江大学出版社) [Zhejiang University Press].
- Li, Dongxiao, and Peter Looms. 2016. "Television Accessibility in China." In *A Multimodal End 2 End Approach to Access*, edited by Pradipta Biswas, Carlos Duarte, Patrick Langdon, and Luis Almeida, 261–273. London: Springer.
- Lin, Zhongxuan, and Liu Yang. 2018. "Denial of Disability Discrimination by Disabled People in China." *Disability & Society* 33 (5): 804–809. doi:10.1080/09687599.2018.1446276.

- Liu, Chih-Ching (刘芷晴). 2015. "Ni Shi Wo de Yan. Yunyong Koushu Yingxiang Fuzhu Guoxiao Shizhangsheng Xuexi zhi Yanjiu" (你是我的眼。运用口述影像辅助國小視障生學習之探究) [You are my eyes: Using audio description to assist the learning of elementary school students with visual impairment]. PhD diss., National Taiwan Normal University.
- Macdonald, Stephen J., Lesley Deacon, Jackie Nixon, Abisope Akintola, Anna Gillingham, Jacqueline Kent, Gillian Ellis, et al. 2018. "The Invisible Enemy: Disability, Loneliness and Isolation." *Disability & Society* 33 (7): 1138–1159. doi:10.1080/09687599.2018.1476224.
- Martínez-Sirés, Paula. 2016. "Making the Inaccessible Accessible: A Panorama of Audiovisual Translation, Accessibility and Audio Description in Japan." *Asiadémica* 7: 24–50.
- Matamala, Anna, and Pilar Orero. 2018. "Standardising Accessibility: Transferring Knowledge to Society." *Journal of Audiovisual Translation* 1 (1): 139–154.
- Orero, Pilar, Stephen Doherty, Jan-Louis Kruger, Anna Matamala, Jan Pedersen, Pablo Romero-Fresco, Sara Rovira-Esteva, et al. 2018. "Conducting Experimental Research in Audiovisual Translation (AVT): A Position Paper." *Jostrans: The Journal of Specialised Translation* 30: 105–126.
- Orero, Pilar, and Irene Tor-Carroggio. 2018. "User Requirements When Designing Learning e-Content: Interaction for All." In *Designing for the User Experience in Learning Systems*, edited by Evangelos Kapros and Maria Koutsombogera, 105–122. Cham: Springer Nature.
- Palmer, James. 2014. "Crippling Injustice." Accessed 14 November 2018. <https://aeon.co/essays/what-is-life-like-for-disabled-people-in-china>
- Pappworth, Maurice Henry. 1967. *Human Guinea Pigs: Experimentation on Man*. London: Routledge and Kegan Paul.
- Rovira-Esteva, Sara. 2010. *Lengua y Escritura Chinas: Mitos y Realidades*. Barcelona: Edicions Bellaterra.
- Snyder, Joel. 2014. *The Visual Made Verbal*. Arlington: American Council of the Blind.
- Tang, Na, and Yu Cao. 2018. "From Multiple Barriers to a Co-Prospersity Society: The Development of a Legal System for Disabled People in China." *Disability & Society* 33, 1170–1174.
- Tor-Carroggio, Irene, and Helena Casas-Tost. (forthcoming). "Who is Currently Audio Describing in China? A Study on the Profile of Chinese Audio Describers." To appear in *MonTl. Monografías de Traducción e Interpretación*.
- Wu, Zongyi (吴宗艺), and Zhenzhen Xie (谢桢桢). 2015. "Zhongguo shizhang koushu yingxiang fuwu de fazhan xianzhuang yu dalu tuiguang" (中国视障口述影像服务的发展现状与大陆推广) [The development and the promotion of audio description in Mainland China]. *Xinwen Yanjiu Daokan* (新闻研究导刊) [Journal of News Research] 6 (10): 20–22.
- Zhang, Chao. 2017. "Nothing about Us without Us': The Emerging Disability Movement and Advocacy in China." *Disability & Society* 32 (7): 1096–1101. doi:10.1080/09687599.2017.1321229.
- Zhou, Xun. 2002. "The Discourse of Disability in Modern China." *Patterns of Prejudice* 36 (1): 104–112. doi:10.1080/003132202128811394.

Annex 2: Additional articles

2.1. Additional article 1

视听翻译新发展：语音合成在口述影像中的应用——基于针对视障人士的接受实验与调查³⁴

肖维青³⁵ Irene Tor-Carroggio³⁶

摘要：口述影像（audio description）作为视听翻译的一种类型，已经为很多视障人士带去了福音，其中语音合成（text-to-speech）口述影像，作为一种高效、经济的手段，已应用于多个语言环境中，如英语、日语、波兰语、西班牙语、加泰罗尼亚语等。鉴于中国亟需推广口述影像服务，本研究通过一项针对上海视障人士的实验与调查，对比语音合成口述影像与真人声音口述影像，探讨了语音合成口述影像在中国的接受。研究结果表明，除了与内容理解相关的指标外，真人声音口述影像的其他指标得分均高于语音合成口述影像。由此可见，中国视障人士更喜欢真人声音口述影像。尽管如此，有超过一半的被测认为，语音合成口述影像可以成为一种过渡甚至永久解决方案，他们期待语音合成口述影像能为视障用户带来更多无障碍电影。

关键词：口述影像；语音合成；接受研究；汉语；无障碍媒体

一、引言

视听翻译（Audiovisual Translation）的研究范畴不仅仅包括我们熟知的配音翻译、字幕翻译、译配解说，还包括为视障人士提供的口述影像以及为听障人

³⁴ Xiao, W. (肖维青), and Tor-Carroggio, I. (2020). Shiting fanyi xin fazhan: yuyin hecheng zai koushu yingxiang zhong de yingyong. Jiyu zhendui shizhang renshi de jieshou shiyan yu diaocha (视听翻译新发展：语音合成在口述影像中的应用——基于针对视障人士的接受实验与调查) [New developments in audio-visual translation: The use of synthetic voices in audio description. An acceptance experiment for people with sight loss]. *Dongfang Fanyi* (东方翻译) [East Journal of Translation], 64(2), 24–30.

³⁵ 上海外国语大学英语学院教授、博士生导师。

³⁶ 西班牙巴塞罗那自治大学与上海外国语大学联合培养博士生，得到孔子新汉学计划资助。

士提供的特殊字幕 (subtitles for the deaf and hard-of-hearing) 等 (Gambier 2003)。近年来, 视听翻译研究把口述影像和特殊字幕归入“无障碍媒体” (media accessibility), 主要关注为残障人士提供的无障碍感官服务。

中国是世界上残障人士总量最多的国家 (吴宗艺、谢桢桢, 2015: 21), 且与日本、英国、美国相似, 也是老龄化国家 (王桦等, 2014: 76)。老龄化社会意味着许多严重问题亟待解决, 包括如何关照残障人士, 因为随着年龄的增长, 罹患残疾障碍的可能性也会上升 (Peng *et al.* 2010)。

自 2007 年正式批准并践行《联合国残疾人权利公约》以来, 中国为提高残障人士的生活质量做出了实质性努力, 近 8300 万残障人士从中受益。《公约》中规定, 残疾人士有权参与文化生活, 获取信息。无障碍感官服务, 例如为视障人士配置口述影像、为听障人士提供字幕或实时手语翻译等, 都可以帮助残疾人实现这一权利。

社会上对特色字幕、实时手语还多少有些了解, 但是对口述影像则知之甚少。按照 Jankowska (2015: 9) 的定义, “口述影像是一种额外添加的语言描述, 以诠释图像上最重要的信息, 是一项帮助视障人士接受视觉信息的技术。”³⁷ 简言之, 就是在戏剧表演、电影电视节目中, 口述影像员描述角色的动作、肢体语言、面部表情、场景以及服装, 通常口述影像插入人物对话的间隔中, 不与重要音效重叠。一般而言, 口述影像者可以是一位专业人士, 也可以是一位经过一定训练的志愿者。

为了促进口述影像的发展, 降低成本, 让更多的视障人士得到口述影像服务, 不少学者开展了相关的接受研究, 调查语音合成的口述影像是否可以取代真人声音的口述影像, 例如 Szarkowska (2011: 145) 认为, “语音合成口述影像的成本效益远高于用传统方法制作口述影像。因为相对提前录制的口述影像, 它不需要录制成本; 而相对实时口述影像, 语音合成不需要人工读稿, 因此可以降低劳动力成本。”

³⁷ 虽然口述影像的大部分研究都针对视觉障碍者, 但有不少研究也强调了口述影像对于移民或语言学习者的潜在益处, 因此其实际运用范围很广, 不仅仅局限于残障人士 (Gambier 2003; Snyder 2014; Walczak 2017; Di Giovanni 2018)。

语音合成口述影像已在加泰罗尼亚语和日语环境下得到测试验证，但尚未在汉语环境中进行任何正式测试，所以无法确定它是否能够提高汉语口述影像的生产速度，从而降低成本。另外，在接受方面，汉语语音合成口述影像的研究结果是否会与其他语种的研究结果有所不同？这两个问题亟待实证研究。

本研究基于建构式复制研究方法，通过调整前人实验中的个别设计，重新评估前人研究结果的信度。研究目的有二：一是通过比较汉语语音合成口述影像与标准真人口述影像在几个关键特征上的异同，评估汉语语音合成口述影像；二是进一步确认在汉语语境下，语音合成口述影像是否能够成为一种过渡或最终解决方案。

二、关于语音合成的视听翻译接受研究

语音合成被广泛应用于不同领域，如全球定位系统、教育平台、娱乐工具等 (Cryer & Home 2008: 5)。其中，语音合成已用于多种形式的视听翻译中，例如旁白 (voice-over)、声音字幕 (audio subtitling)、口述影像。

针对旁白，Matamala 和 Ortiz-Boix (2018) 做过一个对比实验，让 16 位受众比较听到的两段关于野生动物的纪录片，片段时长均在两分钟左右，分别用语音合成与真人配音。总体而言人声得分相对较高，但是在参与度方面，受众均表示两者无明显差异。尽管如此，受众对于未来语音合成代替人声旁白的前景，并不乐观。但是超过一半的受访者表示，他们可以接受电视播放语音合成的声音片段。

语音合成的声音字幕也已经过测试，开始在西班牙、荷兰、丹麦、瑞典等国投入使用 (Iturregui-Gallardo 2019)。早在 2013 年，Thrane (2013) 就评估了 16 位用户在使用语音合成的声音字幕观看新闻、纪录片、非纪实节目中所遇到的问题：同步性、发音、多种声音同时出现、速度等，研究发现，人们更容易接受在新闻等纪实性的节目中使用语音合成。

针对语音合成口述影像的研究，波兰华沙大学的研究人员，就不同体裁的视频素材对语音合成的接受度做了调查，其中包括一系列动画教育片 (Walczak 2010)、一部波兰语故事片 (Szarkowska 2011)、一部波兰语配音的外国电影

(Drożdż-Kubik 2011)、一部配有旁白的外国电影 (Szarkowska & Jankowska 2012)、一部配有声音字幕的纪录片 (Mączyńska 2011)。视频体裁繁多,但是实验结论基本相似:在大部分情况下,语音合成的口述影像都可以作为一种过渡方式被接受,有时甚至是最终方案。相比之下,语音合成口述影像运用于纪实性的视频更加合适。除了波兰语之外,相关研究也涉及了其他欧洲语言,例如西班牙语 (Pazos 2012; Ortiz-Boix 2012)、加泰罗尼亚语 (Fernández-Torné 2016) 等。值得一提的是,虽然中国对于语音合成口述影像研究甚少,但在 2014 年,上海高智科技有限公司利用网上免费的语音合成系统来朗读脚本,为几部电影生产出了相应的口述影像³⁸,但是未经过测试评估,用户反馈调查也未曾进行。

三、研究方法

本研究采用的实验材料由上海市口述影像中心——“光影之声”友情提供。光影之声为非营利性组织,是上海市目前唯一运用数码格式录制口述影像的机构。它每年制作口述影像电影逾 50 部。本研究在分析调查了中国口述影像用户对于电影类型的偏好⁵后,选取了由许鞍华导演的《明月几时有》(2018)作为实验材料,该电影围绕 1940 年代香港抗日运动展开。

为了选取两个有可比性的片段,我们对电影以及口述影像的脚本进行了详细的分析。最终节选了两个视频片段作为实验材料(详见表 1)。两个片段的内容相似,都包含食物的介绍;视频时长均持续五分钟左右;出场人物里两位女主角均有出现;背景均无配乐或其他明显的声音;另外,两个片段的口述影像密集度也保持相对一致,均在 600 字左右。两个片段的开头都描述了方母(主人公的母亲)的举动,并且都提到了镜子和大米。

《明月几时有》原有口述影像使用男声配音,所以我们也选择了男声语音合成,利用国产“科大语音酷 4.0”进行制作。进行混音时,我们得到了上海高智科技有限公司的帮助。虽然语音合成的口述影像可以直接从合成软件中输出,但考虑到研究者需挨家挨户上门至受试者家中,为避免各种问题的出现,研究者采取了最为保险的录制方法。

³⁸ 查询音频可登陆网站 https://www.xfyun.cn/services/online_tts

为了提高语音合成的最终效果，我们还对原口述影像进行了微调，删除了儿化音。因为语音合成无法结合“儿”与前面汉字的结合。所幸儿化音是中国北方方言特有现象，因此即使删除也不会影响理解。

片段	人声配音	语音合成	字数
片段 1	男声	男声	4分56秒 556 字
片段 2	男声	男声	4分37秒 592 字

表 1 所选片段特征

本研究通过问卷调查，评估每个片段中的口述影像声音效果。每位被试需回答两份的问卷，问卷题目由研究者朗读。根据 Chmiel 与 Mazur (2012)，研究者在问卷中加入了理解题。另外列出了 9 个指标，需要被试在听完片段后对这 9 个指标作出评价。其中包括声音自然度、声音愉悦度、停顿、难易程度、理解、发音、语调、接受度和总体印象。³⁹

虽然客观评测是研究人工合成语音的有效途径 (Cryer & Home 2010: 6)，但研究者还是采取了一种相对主观的用户评测方法，因为最终还是由视障用户来使用语音合成的口述影像。另外，“如果想要将语音合成最终运用到产品及服务上，主观用户测评将会变得更加重要，因为可以借此了解用户对此声音是否满意” (Cryer & Home 2010: 4)。实验采取平均意见得分 (Mean Opinion Score)，这是一种用于衡量意见的常用方法。首先请受试者收听语音合成的片段，随后对此从 1 到 10 进行打分，平均大家的分数得出最终得分。

2019 年 6 月，本实验通过了研究者所在大学⁴⁰学术伦理委员会批准，并在正式施行之前针对 3 名用户开展先导实验。实验中，研究者向每位被试朗读实验知情书与同意书，被试口头表示同意参加并录音为证。过程中，每位被试都需欣赏两段视频，分别由语音合成及人声配音，表 2 标明了片段播放顺序，遵循拉丁方阵 (Latin square) 规则，以此往复循环。根据 Fernández-Torné (2016)，实验中均首先播放人工合成语音 (*)，以避免被试对真人声音的好感先入为主。

³⁹ 以上 评判 指标 大部分选取于国际电信联盟 (International Telecommunications Union) 的《标准》(1994: 8) 其中定义了用于 主观 评价合成语音 的方法

⁴⁰ 即上海外国语大学。

被试者	片段一	片段二
01	1*	2
02	2*	1

表 2 播放顺序

随后研究者在 Szarkowska、Jankowska (2012) 和 Fernández-Torné (2016) 研究成果的基础上制定了最终问卷。问卷的语言，简洁易懂，因为许多被测可能为老年人，文化程度偏低。

在上海导盲犬组织与上海残疾人联合会的协助下，研究者邀请到了 40 位视障被试进行实验，对被试的唯一的的要求是之前曾接触过无障碍电影即口述影像。实验过程均在上海展开，地点包括被测家中、社区中心、华东师范大学。在华师大进行的实验还包括一项 10 人的焦点小组实验，10 名被测分享了自己对于语音合成的口述影像与中国口述影像的看法。

最后，本研究使用 SPSS25.0 版完成实验结果的定量分析。由于样本数量限制，我们采取了非参数检验的方法，例如威尔科克森符号秩检验 (Wilcoxon signed-rank test) 和斯皮尔曼等级相关系数 (Spearman's rank correlation coefficient)。P 值需要达到 0.05，才可以说数据之间具备了显著性差异。

四、 结果与讨论

我们从样本信息入手，然后根据定量与定性数据，集中讨论本实验结果，

(一) 样本描述

本次研究共有 40 名口述影像使用者参加，其中男性占比 52.5%，女性占比 45%，（1 位被试没有回答这个问题）。与预期相似，大部分被测年龄较大，平均年龄为 58.64 岁，中位年龄为 61 岁，其中年龄最小的为 30 岁，最大的为 78 岁。在教育程度方面，大部分人（72.5%）为中学或职业学院毕业，还有约 17.5% 的被试具有大学学历。另外，1 名被试为小学毕业，1 名被试没有接受过教育，还有 2 名被试拒绝回答这个问题。在语音合成的使用方面，有 85% 的被试称之前曾接触过（见图 1），大部分是在手机或者电脑上，还有超过一半的被试称经常或者一直使用语音合成（见图 2）。

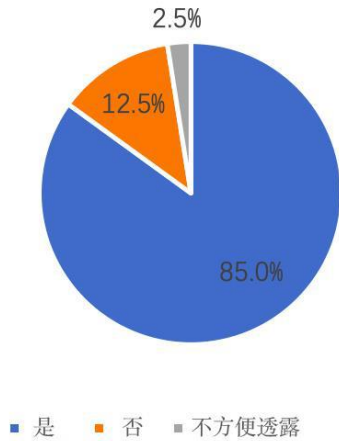


图1 是否接触过语音合成

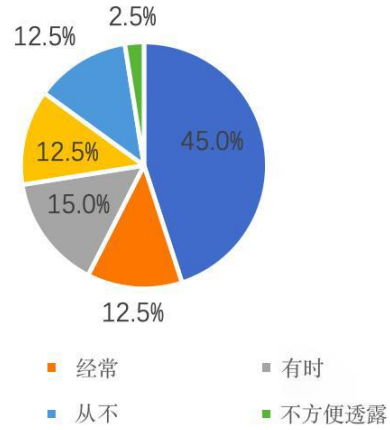


图2 使用语音合成的频率

与实验要求一致，所有被试均至少有过一次口述影像体验，其中超过一半的被试称经常或一直使用口述影像（67.5%）。

（二）内容的理解与声音的质量

在听完两则视频片段后，被试需要回答3个关于内容理解的问题。令人惊讶的是，被试在听完语音合成与真人声音的口述影像后，平均答对题数分别只有1.42与1.75，且没有显著差异（ $Z = -1.62$, $p = 0.10$ ）。这一结果可能是多方面因素造成的，比如被试年龄偏大；比如实验前没有向被试提及要进行理解测试，所以被试没有刻意关注内容细节。之所以这样设计，是考虑到测试的生态效度，希望被试不会为了完成测试而过分注意每个细节。另外，几乎一半的实验都是在被试家中进行，无法保证周围环境完全安静。而大部分被试不便离家，所以在大多数情况下，在家中实验是研究者唯一的选择。当然，这样也可以借机在更真实的环境下观察口述影像用户的反应，毕竟语音合成的口述影像最终还是在类似的环境下使用的。

为了研究口述影像用户对内容理解的影响因素，我们决定引入一组年轻被试，让他们在类似条件下完成口述影像电影片段的欣赏。50名视力正常的大学生同意参加实验，他们是浙江外国语学院翻译专业大三的学生。实验在他们的教室中进行，实验期间他们均被蒙住双眼。本组实验中，在欣赏过语音合成及真人声音的口述影像后，学生们平均答对题数分别为1.80与2.10。虽然结果显示这批被试可以更好地理解真人声音的口述影像，但在统计意义上没有显示出差异（ $Z = -1.81$, $p = 0.07$ ）。学生组年龄较小，基本都在20岁左右，所以成绩相对更好。另外他们的

教育水平更高，教室作为实验场所也相对更加安静。但是在统计数据上，无论是针对语音合成 ($Z = -1.92$, $p = 0.054$) 口述影像片段还是真人声音口述影像片段 ($Z = -1.62$, $p = 0.10$)，两组被试的结果均没有显著差异。

虽然我们引入了学生组实验中，但本文重点分析视障人士被测组的数据，因为他们才是口述影像的主要受惠者。因此以下数据均来源于视障人士被测组的研究结果。

表 3 中包括两种声音 9 项指标的平均意见得分，用于比较声音的质量。

指标	语音合成平均意见得分	人声平均意见得分
声音自然度	8.35	9.60
声音愉悦度	8.20	9.18
语调	7.65	9.25
发音	9.10	9.60
停顿	8.77	9.38
难易程度	9.20	9.63
理解性问题	9.80	9.85
接受度	9.13	9.68
总体印象	8.57	9.33

表 3 描述性结果

两种声音均在理解方面得分最高，并且所有的平均得分均在 7.5 分之上。真人声音的各项得分在 9 分以上，比语音合成要高。但是语音合成也有多项得分超过了 9 分，例如发音、难易程度、理解、接受度。与预期相似，得分最不理想的是语调，研究人员 (Walczak 2010) 曾经指出，语音合成主要的缺点之一就是不自然的语调。

表 4 显示，在几乎所有的指标中，两种声音均有统计差异。因此可以肯定地说，真人声音的质量高于语音合成。但是在理解方面，二者没有显示出明显的差异。

指标	威尔科克森符号秩检验结果
声音自然度	$Z = -3.29$

	$p < 0.01$ ties = 10
声音愉悦度	$Z = -2.73$ $p < 0.01$ ties = 11
语调	$Z = -3.51$ $p < 0.01$ ties = 11
发音	$Z = -2.29$ $p = 0.02$ ties = 21
停顿	$Z = -2.40$ $p = 0.01$ ties = 21
难易程度	$Z = -1.99$ $p = 0.05$ ties = 22
理解性问题	$Z = -1.42$ $p = 0.15$ ties = 32
接受度	$Z = -2.04$ $p = 0.04$ ties = 20
总体印象	$Z = -2.60$ $p < 0.01$

表 4 威尔科克森符号秩检验结果

(三) 被试的偏好

在定量数据基础上，被试还针对使用口述影像的偏好与语音合成的运用，提供了一些定性评价。大部分被试表示，只要最终可以与视力正常的人一起欣赏电影，他们可以接受任何形式，包括使用耳机。

当被问及更喜欢哪种声音的时候，一半的被试选择了真人声音，但是也有些人对真人声音不太满意，因为他们认为真人有口音，而且听上去更像是在背诵而非描述。另外有 27.5% 的被测表示没有明显的偏好，两种声音都可以接受，还有 20% 的被测更喜欢语音合成的声音（图 3）。预料之外的是，人工合成的声音因为十分清晰，而受到了大家的认可，并且有些被试称合成声音十分平稳，没有口音，因此更加喜欢语音合成。Szarkowska 和 Jankowska (2012: 85) 也发现许多用户喜欢不带口音的声音。

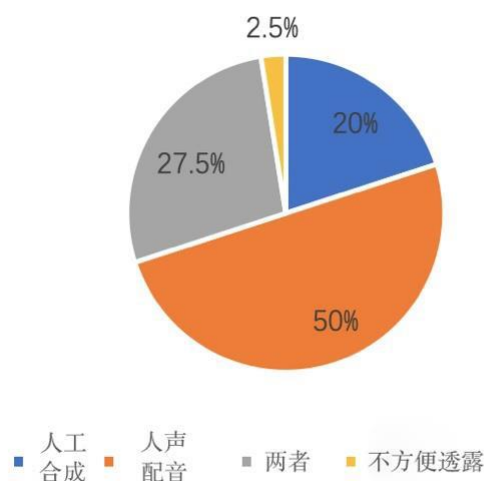


图 3 声音的偏好

当问及更想要哪种声音用于口述影像的电影中时，超过一半的人选择了真人声音（57.5%），而只有一人选择语音合成（2.5%）。虽然差距巨大，但是有 15% 的被测表示无所谓，还有将近 1/4 的被测认为这取决于电影类型。这些被测认为 11 纪录片的口述影像更适合纪录片的口述影像更适合使用语音合成用语音合成⁴¹，然而然而纪录片并不是运用口述影像最广纪录片并不是运用口述影像最广、受益受益面最大的电影类型最大的电影类型。此外此外，被试被试认为语音合成也认为语音合成也更适用于更适用于历史及军事历史及军事题

⁴¹ 这与 Fernández Torné 的发现相符 (2016)。

材题材作作品。一位一位被测特别被测特别强调，口述影像是一门艺术，所以强调，口述影像是一门艺术，所以语音合成的口述影像并不适语音合成的口述影像并不适用于所有类型的电影。但是用于所有类型的电影。但是所有被测都一致认为所有被测都一致认为，口述影像口述影像的内容的内容本身比所用的本身比所用的声音更重要。声音更重要。

虽然量化与质化数据都显示被测数据都显示被测明显更偏好明显更偏好真人声音真人声音，但除了，但除了 11 位位被测被测之外之外，，其他其他被测均被测均表示可以接受语音合成作为一种过渡方案（图表示可以接受语音合成作为一种过渡方案（图 44），还有将近），还有将近 63%63%的的被测被测甚至可以接受它作为最终的甚至可以接受它作为最终的解决解决办法（图办法（图 5））。

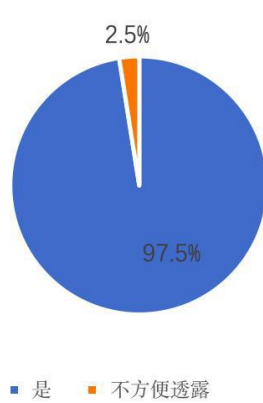


图 4 过渡方案接受度

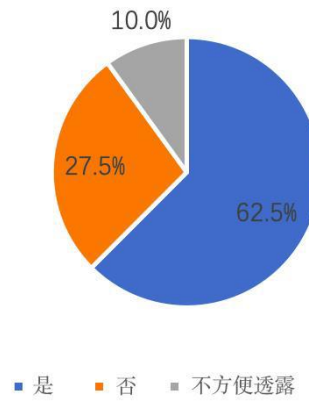


图 5 最终方案接受度

如表 5 所示，本研究数据与前人研究结果相似。

研究	语音/视听样本	过渡方案接受度	最终方案接受度
Szarkowska (2011)	波兰语/波兰语故事片	95%	58%
Szarkowska & Jankowska (2012)	波兰语/配有波兰语旁白的外国电影	95%	70%

Fernández-Torné (2016)	加泰罗尼亚语/各种类型的电影	94%的人认为语音合成口述影像可以作为真人声音的替代	
本研究	汉语/历史类电影	97.5%	62.5%

表 5 与同类研究的比较

在研究语音合成口述影像在中国的接受度时，必须要考虑到中国口述影像电影匮乏的问题，这导致了用户的选择余地很小。许多被试抱怨没有足够的电影可以选择，特别是没有新电影和外国电影。Szarkowska 与 Jankowska (2012: 84) 指出，视障人士希望能像视力正常的人一样欣赏外国电影。所以如果有办法可以帮助他们接触到更多的无障碍电影，他们都抱着开放接受的态度。尽管如此，并不是所有人都是因为口述影像稀缺而接受语音合成。有些被试无法区分真人声音与合成声音，还有人甚至没有意识到本测试中第一个电影片段是由语音合成完成的。有被试表示，即使是真人声音也不能保证口述影像的效果。有些被试承认每个月在影院欣赏完无障碍电影之后，他们会去确认实时口述影像的配音人员，看看自己是否喜欢这位志愿者⁴²。也有被试人指出，有时候很难区分电影对话部分与口述影像部分，因为它们的配音可能会十分相似。使用语音合成恰恰可以解决这一问题。

还有两位被试提出，语音合成技术正在不断进步，所以他们相信在不久的将来，语音合成在口述影像中将会代替传统人声配音，另外还有被试想了解克隆人声的语音合成技术。上述两位被试也表示，如果口述影像服务实行收费可以实现更多的影片选择，提供线上资源，他们乐意支付这笔费用，这样他们就可以在家享受电影。他们还重点强调，由于口述影像服务的志愿性质，用户必须时刻心怀感激，且不能及时指出口述影像中存在的问题，这些都对口述影像服务的使用体验产生了消极影响。

然而，语音合成口述影像最大的问题在于声音过于单调，不含任何情绪。不过相较于口述影像资源的稀缺问题，这个问题也无伤大雅了。

⁴² 上海有一部分广播电视节目主持人会担任志愿者，自发为视障人士提供口述影像服务。

最后，我们必须研究语音合成的接受度与用户使用语音合成或口述影像的频率之间是否相关。分析得出，两者均并没有呈现统计意义上的显著联系 ($r = 0.06$, $p = 0.72$ / $r = 0.08$, $p = 0.61$)。

五、 结语

本研究通过与传统口述影像对比，探索了语音合成口述影像在中国的接受。研究结果显示，真人声音口述影像得分整体高于语音合成口述影像，视觉障碍者更喜欢真人声音的口述影像。当然，这并不意味着语音合成口述影像不被接受，因为超过一半的被测将其视为一种过渡的替代方案，甚至是一种最终解决办法。

虽然之前已有针对语音合成口述影像的用户测试，但本研究在以下几方面具有创新性。第一，语音合成测试尚未在汉语环境下开展。在中国大陆，从未有过针对口述影像的实验。因此本研究希望能够激发其他研究者，特别是对无障碍媒体，尤其是用户测试研究的兴趣。第二，相对于 Fernández-Torné (2016) 的研究，本实验中采用的两个视频片段时长更久。这也是参考了 Fernández-Torné 的建议，有助于将此类研究向前推进。第三，本研究以用户为中心，所选实验材料也是基于实验前的用户调查结果。第四，本实验在问卷中增加了内容理解题，且将这部分问题安排在声音评价问题之前。Walczak (2010) 曾在针对视障儿童的实验中加入了内容理解题，但这一做法目前并未普及。本实验如此设计旨在于测试被测对内容究竟理解了多少，因为口述影像用户对于内容的理解往往被忽视，但实际上这一点甚至要比技术创新更重要。

本研究当然还存在着一些局限性。第一，实验中只采用了一种语音合成的声音与一种真人声音进行比较。两种声音都不是由口述影像用户选择的，所以我们无法确定如果由他们选定声音，实验结果是否会有所不同。第二，中国幅员辽阔，社会经济发展不平衡，本实验仅在上海一地开展，因此实验结果并不一定能代表整个国家的情况。第三，实验样本偏小，且非随机抽样，这一点需在未来研究中改进。尽管如此，笔者还是希望抛砖引玉，对未来的研究有所启发。

参考文献

- Chmiel, Agnieszka and Iwona Mazur (2012). “AD reception research: some methodological considerations”. Elisa Perego (ed.) (2012) *Emerging topics in translation: audio description*. Trieste: EUT, 57-74.
- Cryer, Heather and Sarah Home (2008). Exploring the use of synthetic speech by blind and partially sighted people. Literature review #2. Birmingham: RNIB Centre for Accessible Information (CAI).
- Di Giovanni, Elena (2018). “Audio Description for Live Performances and Audience Participation”. *JoSTrans: the Journal of Specialised Translation* 29, 189–211.
- Drożdż-Kubik, Jan (2011). *Harry Potter i Kamień Filozoficzny słowem malowany – czyli badanie odbioru filmu z audiodeskrypcją z syntezą mowy*. Masters dissertation. Jagiellonian University.
- Fernández-Torné, Anna (2016). *Audio Description and Technologies. Study on the Semi-Automatisation of the Translation and Voicing of Audio Descriptions*. PhD thesis. Autonomous University of Barcelona.
- Gambier, Yves (2003). “Introduction”. *The Translator*, 9(2), 171–189.
- Iturregui-Gallardo, Gonzalo (2019). *Audio Subtitling: Strategies for Voicing Subtitles and Their Effect on Film Enjoyment and Emotion*. PhD thesis. Autonomous University of Barcelona.
- Jankowska, Anna (2015). *Translating Audio Description Scripts: Translation as a New Strategy of Creating Audio Description*. Berna: Peter Lang.
- Matamala, Anna and Carla Ortiz-Boix (2018). “Text-to-speech voice-over? A study on user preferences in the voicing of wildlife documentaries.” *Skase. Journal of Translation and Interpretation* 1, 24-38.
- Mączyńska, Magdalena (2011). *TTS AD with audio subtitling to a non-fiction film. A case study based on La Soufriere by Werner Herzog*. Masters dissertation. University of Warsaw.

Ortiz-Boix, Carla (2012). *Tecnologies per a l'audiodescripció. Estudi sobre l'aplicació de la traducció automàtica i la síntesi de parla a l'audiodescripció en castellà*. Masters dissertation. Autonomous University of Barcelona.

Pazos, Patricia (2012). *Audiosubtitulació: una posible solució para la accesibilidad a los medios audiovisuales*. Masters dissertation. Autonomous University of Barcelona.

Peng, Xiaoxia, Shige Song, Sheena Sullivan, Jingjun Qiu and Wei Wang (2010). "Ageing, the urban-rural gap and disability trends: 19 years of experience in China -1987 to 2006." *PLoS ONE*5(8), e12129.

Snyder, Joel (2014). *The Visual Made Verbal: A Comprehensive Training Manual and Guide to the History and Applications of Audio Description*. Arlington: American Council of the Blind. Abingdon: Routledge.

Szarkowska, Agnieszka (2011). "Text-to-speech audio description: towards wider availability of AD." *Journal of Specialised Translation*15, 142-163.

Szarkowska, Agnieszka, and Anna Jankowska (2012). "Text-to-speech audio description of voice-over films. A case study of audio described *Volver* in Polish." Elisa Perego (ed.) (2012). *Emerging Topics in Translation: Audio description*. Trieste, Italy: EUT, 81-98.

Thrane, Lisbeth (2013). *Text-to-speech on Digital TV An Exploratory Study of Spoken Subtitles on DRISyn*. Masters dissertation. University of Copenhagen.

Tor-Carroggio, Irene, and Helena Casas-Tost (submitted). "Who Is Currently Audio Describing in China? A Study of Chinese Audio Describer Profiles".

Tor-Carroggio, Irene (in press). "The customer is always right: study on Chinese persons with sight loss' opinion on their experience with audio description". *Disability & Society*. doi: 10.1080/09687599.2020.1713727

Tor-Carroggio, Irene and Gert Vercauteren (2020). "When East Meets West: A Comparison of Audio Description Guidelines in China and Europe". To appear in *HIKMA. Translation Studies Journal*.

Walczak, Agnieszka (2010). *Audio description for children. A case study of text-to-speech audio description of educational animation series Once Upon a Time... Life*. Masters dissertation. University of Warsaw.

Walczak, Agnieszka (2017). "Audio Description on Smartphones: Making Cinema Accessible for Visually Impaired Audiences". *Universal Access in the Information Society*17(4), 833-840.

王桦, 赵晟珣, 曾尔亢, 马春薇, 王家瑜, 段凌. 中国人口老龄化社会发展与应对策略. 中国社会医学杂志, 2014, 第 31 卷(2): 75-77.

吴宗艺, 谢桢桢. 中国视障口述影像服务的发展现状与大陆推广. 新闻研究导刊, 2015, 第 6 卷(10): 20-22.

2.2. Additional article 2

Comer, beber, amar...y comparar. Análisis contrastivo de la audiodescripción en chino y español: un estudio de caso⁴³

Resumen

El presente artículo lleva a cabo un análisis contrastivo de una audiodescripción en chino y en español para dirimir si las diferencias entre estas dos lenguas y culturas se traducen en guiones de audiodescripción significativamente diferentes. Nuestro estudio de caso se vertebra alrededor de la película *Comer, beber, amar* de Ang Lee (1994) y su análisis se divide en dos partes: en primer lugar, detectamos las diferencias en las descripciones de los personajes; y, en segundo lugar, estudiamos los culturemas, que categorizamos de acuerdo con la propuesta de Molina (2001). A continuación, analizamos las técnicas de traducción empleadas para su traslado partiendo de la clasificación de las técnicas traductoras propuesta por Molina y Hurtado (2002). Los resultados demuestran que existen numerosas diferencias en la manera de audiodescribir entre el par de lenguas escogido, hecho que afecta directamente a la traducción de guiones de audiodescripción, una manera más rápida y barata que podría emplearse en el futuro en China para ofrecer este servicio de accesibilidad.

Abstract

This article conducts a contrastive analysis of an audio description in Chinese and Spanish to decide whether the differences between these two languages and cultures translate into significantly different audio description scripts. Our case study is based on the movie *Eat, Drink, Men, Woman* by Ang Lee (1994) and its analysis is divided into two parts. First, we detected the differences in the descriptions of the characters; and, second, we studied the culturemes, which we categorized according to Molina's (2001) classification. Then we analyzed the translation techniques used to transfer the culturemes using Molina and Hurtado's (2002) categorization. The results show that there are numerous differences in

⁴³ Liu, Y., and Tor-Carroggio, I. (2022). *Comer, beber, amar... y comparar. Análisis contrastivo de la audiodescripción en chino y español: un estudio de caso [Eat Drink Man Woman... A Contrastive Analysis of an AD in Chinese and in Spanish: A Case Study]*. *Onomázein*.

the way audio description is created between the chosen language pair. This will have an impact on the translation of audio description scripts, a faster and cheaper way to create audio description, which could be used in the future to offer this access service in China.

Palabras clave: audiodescripción, chino, técnicas de traducción, culturema, análisis contrastivo

Keywords: audio description, Chinese, translation techniques, cultureme, contrastive analysis

1. Introducción

Pese a que China es el país con más personas con diversidad funcional del mundo (Wu y Xie, 2015), en comparación con muchos países europeos, la accesibilidad a los medios es todavía una de sus grandes asignaturas pendientes. China ratificó la Convención sobre los Derechos de las Personas con Discapacidad de las Naciones Unidas en 2008, pero servicios de accesibilidad sensorial tales como la subtitulación para personas sordas y la audiodescripción (AD) se encuentran todavía en un estado muy incipiente. De hecho, el concepto de AD es aún muy desconocido para el gran público chino (Wu y Xie, 2015) y, a diferencia de países vecinos como Japón y Corea del Sur —donde la AD se ofrece en la televisión nacional— (Leung, 2018), en China este servicio de accesibilidad no se encuentra disponible en la totalidad del país y su ámbito de uso es aún muy restringido, siendo las películas su principal aplicación (Tor-Carroggio y Casas-Tost, próximamente). Además, las ciudades que sí ofrecen este servicio lo hacen, sobre todo, de forma gratuita y gracias a la ayuda de voluntarios, que escriben los guiones de AD (GAD) y los locutan, ya sea en directo o bien para su posterior grabación (Tor-Carroggio y Casas-Tost, próximamente). Asimismo, pese a que existen varios grupos de voluntarios de AD en todo el país, no existen guías ni estándares unificados que faciliten la oferta de un servicio homogéneo y de calidad a lo largo y ancho del país (Tor-Carroggio y Vercauteren, próximamente). Del mismo modo, todavía no se ha legislado en el ámbito de la AD, hecho que imposibilita el reclamo de su oferta y niega la posibilidad de amparo legal a las personas que requieren este servicio de accesibilidad.

En vista de que uno de los problemas a los que la AD en China debe hacer frente en la actualidad es la escasez de voluntarios que redacten los GADs (Tor-Carroggio, próximamente), nos proponemos investigar una nueva forma de crear AD de forma más barata y rápida en China: la traducción de GADs ya existentes en otros idiomas, más concretamente del español al chino. Este método ya ha sido explorado e incluso testado con otras combinaciones lingüísticas mediante estudios de recepción en otros países y ha obtenido resultados positivos (López Vera, 2006; Jankowska, 2015). Antes de poner a prueba empíricamente nuestra propuesta con usuarios, es necesario estudiar su viabilidad, por lo cual es preciso detectar y radiografiar las similitudes y diferencias a la hora de audiodescribir en España y China. Este es, precisamente, el objetivo del presente artículo: analizar, mediante un estudio de caso, una AD disponible tanto en chino como en español (*Comer, beber, amar*, de Ang Lee, 1994) centrándonos en dos aspectos relevantes para la interpretación de la película: las descripciones de los personajes y el traslado de los culturemas. Más concretamente, en el presente artículo nos hemos planteado los siguientes tres objetivos:

- Sistematizar las diferencias en las descripciones de los personajes y de los culturemas en dos ADs del mismo original.
- Explicar las razones que pueden motivar las diferencias en las descripciones de los personajes y de los culturemas en cada una de las versiones.
- Analizar las técnicas utilizadas en las dos ADs en el caso de los culturemas.

En relación con los susodichos objetivos, hemos partido de las siguientes dos hipótesis: en primer lugar, debido a las diferencias lingüísticas y socioculturales entre el par de lenguas estudiado, se espera que existan diferencias significativas tanto en las descripciones de los personajes como en las de los culturemas. En segundo lugar, se espera que las diferencias socioculturales afecten a las técnicas utilizadas para trasladar los culturemas.

Este artículo se divide en cuatro secciones. Primero efectuaremos una revisión de los estudios de caso que contrastan la AD en diferentes idiomas, incluidos aquellos que aportan información acerca de la AD en China. También repasaremos algunos que tratan los culturemas. A continuación, describiremos la metodología empleada en nuestro estudio. Luego, presentaremos los resultados principales obtenidos y unos errores encontrados en el GAD español, finalmente, expondremos nuestras conclusiones y las

limitaciones a las que nuestra investigación ha tenido que hacer frente, así como algunas líneas de investigación que podrían explorarse en el futuro.

2. Antecedentes de investigación

Aunque la AD ya es una realidad en muchos países, resulta lógico que, a causa de las diferencias socioculturales y lingüísticas, las técnicas de AD difieran entre países distintos. Existen numerosos estudios que contrastan la AD en diferentes idiomas. Bourne y Jiménez Hurtado (2007) compararon una AD en español e inglés (*The Hours*, de Stephen Daldry, 2002) para explorar la viabilidad de la traducción de GADs en este par de lenguas. Estos investigadores identificaron bastantes diferencias tanto en el contenido como en la forma de la AD, ya que se percataron, por ejemplo, de que en el GAD inglés se usaban más oraciones simples mientras que en el español abundaban las oraciones complejas. Aparte de las diferencias lingüísticas y estilísticas, Matamala y Rami (2009) estudiaron la manera de tratar los referentes culturales en una película alemana audiodescrita tanto en español y como en alemán (*Good-bye Lenin*, de Wolfgang Becker, 2003). Sus resultados apuntan que las diferencias culturales desembocan en técnicas distintas a la hora de audiodescribir los elementos culturales. Limbach (2012) realizó un estudio de caso basándose en una película audiodescrita en alemán y en inglés (*Slumdog Millionaire*, de Danny Boyle, 2008) con el fin de descubrir cómo afecta el grado de neutralidad de las descripciones a la comprensión de los receptores. Arma (2016) llevó a cabo un estudio de caso entre el italiano y el inglés centrándose en el registro y en la selección de elementos visuales de la película *Chocolat* (Lasse Hallström, 2000). Arma (2016) descubrió que las diferencias más relevantes residen en la manera de describir los personajes, sobre todo, su aspecto físico y sus emociones. En el proyecto VIW (2016) se creó un corpus de acceso abierto que alberga 30 versiones de AD de la misma película en español, catalán e inglés. Además, también contiene ADs realizadas por estudiantes de máster formados en AD. Matamala (2018) comparó las susodichas ADs para determinar si existen similitudes y diferencias relevantes entre las ADs producidas por profesionales y alumnos con el fin de mejorar la formación de los audiodescriptores.

Por otro lado, los elementos culturales siempre han recibido mucha atención en el mundo de la traductología, hecho que se refleja en los múltiples términos que existen para referirse a ellos: “palabras culturales extranjeras” (Newmark, 1992 [1988]), “realia”

(Vlashov y Florin, 1970), “indicadores culturales o puntos ricos” (Nord, 1994) o “referentes culturales” (Mayoral, 1994; Santamaría, 2001b). En nuestro artículo hemos optado por hablar de “culturemas”, una propuesta de Molina (2001: 89) basada en Nord (1997: 34): “Entendemos por culturema un elemento verbal o paraverbal que posee una carga cultural específica en una cultura y que, al ser transferido a otra cultura, puede provocar una transferencia nula o distinta al original”. Pese a que, debido a la distancia cultural, los culturemas sean donde posiblemente más puedan discrepar los GADs en lenguas diferentes, se trata de un tema poco estudiado en el ámbito de la AD (Jankowska y otros, 2017).

En vista de los estudios mencionados, podemos concluir que la comparación de GADs en diferentes idiomas se ha llevado a cabo entre idiomas cercanos hasta la fecha y que el tratamiento de los culturemas en la AD no está muy estudiado, por lo que nuestra propuesta supondría dar un paso más allá. Además, cabe destacar que, como ya hemos indicado, la práctica de la AD aún no se encuentra muy extendida en China, por lo que no es de extrañar que los estudios académicos que versan sobre ella sean bastante escasos (Chao, 2002; Yeung 2007; Leung, 2018; Li, 2013; Tor-Carroggio y Casas-Tost, próximamente). Por tanto, consideramos que este estudio, aparte de proponer una nueva combinación lingüística, constituye una aportación al, prácticamente virgen, estudio de la AD en china. Una aportación que, por otro lado, puede tener un impacto social muy relevante en un futuro cercano.

3. Metodología

Antes de seleccionar la película objeto de estudio, se establecieron unos criterios de selección que se dirimieron imprescindibles. En primer lugar, la película debía tener AD tanto en chino como en español. En segundo lugar, era necesario que la película, en caso de que no estuviera filmada originalmente ni en chino ni en español, estuviera doblada en estos dos idiomas para así poder prescindir de la audiosubtitulación, que no era nuestro foco de estudio. En tercer lugar, la película debía contener gran cantidad de culturemas, puesto que, tal y como hemos detallado, el análisis de estos constituye una parte importante de la investigación que planteamos. Optamos finalmente por una película de Ang Lee: *Comer, beber, amar* (1994), puesto que cumplía con todos los requisitos.

La trama de la película tiene lugar en Taipei. Chu es un chef jubilado y viudo, que vive en una casa vieja con sus tres hijas, con las que no guarda una buena relación. Pese a ello, cenar juntos todos los domingos, reunión que simboliza la unidad familiar. La hija mayor, Jen, es una profesora muy conservadora y rígida. La segunda hija, Chien, tiene talento para la cocina, pero su padre le prohibió seguir sus pasos y ahora es una ejecutiva de una aerolínea. La hija menor, Jia-Ning, todavía está cursando sus estudios y trabaja a tiempo parcial en una hamburguesería. Jia-Ning y Jen abandonan el hogar, descomponiendo así el núcleo familiar. Primero se marcha Jia-Ning, para mudarse con su novio. Luego se va Jen para casarse. Paralelamente muere el mejor amigo de Chu y llega la señora Liang, madre de Jin-Rong. Todas las decisiones importantes se anuncian en las cenas de los domingos. Finalmente, Chu invita a la familia de Jin-Rong a una cena dominical, durante la cual Chu anuncia que ha decidido vender la casa vieja y casarse con Jin-Rong, una muchacha mucho más joven que él.

Para llevar a cabo nuestro análisis de la AD de la misma película, primero transcribimos ambos GADs. A continuación, clasificamos manualmente las descripciones de los personajes en función de si versaban sobre su aspecto físico o su identificación. Después categorizamos los elementos culturales encontrados en los dos guiones y las técnicas usadas para su traslado aplicando la clasificación de los ámbitos culturales de Molina (2001) y la de las técnicas de traducción propuesta por Molina y Hurtado (2002). La clasificación de los elementos analizados se llevó a cabo con el programa ATLAS.ti.

En los ejemplos que se presentan para ilustrar los hallazgos se han puesto los elementos que se desean destacar en cursiva en el GAD español y se han dejado sombreados para el caso del chino. Además, los adjetivos abstractos se han subrayado para destacar su presencia.

4. Análisis contrastivo de la AD de *Comer, beber, amar*

Esta sección presenta y comenta los resultados de nuestro análisis, que gira alrededor de dos grandes bloques, a saber, la descripción de los personajes y el traslado de los culturemas.

4.1. Análisis de las descripciones de los personajes

Los personajes y sus acciones forman parte importante de la narrativa fílmica, pero, a causa de la limitación temporal a la que se enfrenta la AD, es necesaria una priorización de los elementos visuales en función de su relevancia. En este apartado, haremos una comparación de las descripciones físicas de los protagonistas y estudiaremos la forma de identificarlos.

4.1.1. Descripciones físicas de los protagonistas

A partir de nuestro análisis, y tal y como hemos plasmado en la Tabla 1, hemos observado que, para describir la apariencia física de una persona, en primer lugar, se usan más adjetivos en el guion español que en el chino. También hemos descubierto que se utilizan más adjetivos abstractos en el GAD español.

Tabla 1 Adjetivos usados en las descripciones físicas de los personajes

Personaje	AD en español	AD en chino
Chu	sexagenario, entrecano	
Jen	<u>delicada</u> , <u>oriental</u> , bien dibujados , <u>tristes</u> , <u>melancólicos</u> , maquillada, suelto, rizado	橘红色的 (anaranjado)
Chien		短 (corto), 时尚的 (de moda), 白领 (de oficinista)
Jia-Ning	joven, <u>rebelde</u>	挎着单肩包的 (con una bandolera)
Jin-Rong	corto, <u>dulces</u>	
Li Kai	<u>guaperas</u>	高大的 (alto), <u>清秀的</u> (refinado)
Ming-Dao	negra	黑 (negras)
Señora Liang	madura	<u>妖艳</u> (de manera seductora)
Wen		红色的 (rojo)

Cabe destacar que el uso de adjetivos abstractos muchas veces indica directamente un rasgo de la personalidad de un personaje. En el siguiente ejemplo podemos observar la presentación de Jen en el GAD español:

Ejemplo 1: Jen es una delicada belleza oriental con cara de óvalo y labios bien dibujados. Aunque sus ojos son tristes y melancólicos.

Para referirse a la apariencia física de Jen, se han usado palabras como delicada, oriental, tristes y melancólicos. Todo esto fomenta la construcción de la imagen de una mujer sentimental, triste, exótica e inalcanzable. Cabe mencionar el uso del adjetivo oriental en el GAD español, que aparece en más de una ocasión tanto para referirse a personas como a objetos. Teniendo en cuenta que se trata de una película china, no resulta reseñable que todos los personajes sean orientales, por lo que el uso de esta palabra puede confundir al espectador ciego, quien puede creer que, exceptuando los casos en los que se especifica el origen, el resto de los personajes son occidentales. Por ello, consideramos este detalle impreciso y redundante. Por otro lado, el uso de adjetivos abstractos priva a los espectadores ciegos de la opción de configurarse su propia impresión general del personaje por sí mismos.

Asimismo, hemos observado que en un gran número de ocasiones los dos guiones no coinciden en la elección de las características físicas que se van a describir. Por ejemplo, en la primera descripción en el GAD español de Ming-Dao, el novio de Jen, se mencionan sus tejanos y su cazadora negra. Por el contrario, en el chino se opta por describir sus gafas de sol negras (Tabla 2):

Tabla 2 Descripción de Ming-Dao

<p>Descripción en español: Ming-Dao se aleja en su moto. Viste <i>tejanos</i> y <i>una cazadora negra</i> con <i>letrero</i> en la espalda.</p>	<p>Descripción en chino: 校门口，戴着黑墨镜的男教练，骑着摩托车，朝家珍这边驶过来。</p> <hr/> <p>Traducción de la descripción en chino: En la puerta del instituto, el entrenador con <i>gafas de sol negras</i>, se dirige hacia Jen en una motocicleta.</p>
--	--

Pese a estas diferencias, sí existen algunos puntos en común en ambas versiones como, por ejemplo, la escasez, e incluso inexistencia, de descripciones físicas de los personajes.

Por ejemplo, en el GAD español no hay descripción de la apariencia de Chien y en el chino no la hay para Chu a pesar de la gran importancia de los dos papeles.

Otra similitud destacable es que muchas veces en los dos GADs se usa un adjetivo abstracto para resumir la impresión general que confiere un personaje en lugar de proporcionar una descripción más detallada. Por ejemplo, al hablar de Li Kai, nuevo compañero de trabajo de Chien, en el GAD español se dice directamente que es muy guapo sin una descripción de sus características físicas:

Ejemplo 2: Ejecutivo *guaperas*.

Y para describir a la señora Liang, el GAD chino dice:

Ejemplo 3: 锦荣妈妈打扮得很妖艳。(La madre de Jin-Rong se viste de manera *seductora*.)

Aunque algunas características de personalidad se pueden inferir a través de, por ejemplo, los diálogos y la voz de cada personaje, creemos que en este caso concreto harían falta más descripciones físicas, ya que la apariencia puede ayudar al público ciego a caracterizar los personajes de manera más precisa. Según la teoría de esquemas de Bartlett (1932), los espectadores pueden forjarse una representación mental del patrón de comportamientos y pensamientos de un tipo de personas basándose en su apariencia física. Por lo tanto, las descripciones físicas nos ayudan a clasificar a las personas, hecho que conllevará la creación de unas expectativas acerca de su personalidad y de sus acciones posteriores. Dichas expectativas son las que permitirán que la AD sea más sucinta y, al mismo tiempo, aliviarán así la carga cognitiva de los espectadores.

4.1.2. Identificación de los personajes

En el GAD español se nombra al personaje después de su aparición mientras que en el chino se hace la primera vez que aparece en escena (Tabla 3).

Tabla 3 Identificación de Chu en su primera aparición

Descripción en español:	Descripción en chino: 镜头转到了老朱家的院子。
Dentro <i>un hombre</i> agarra un pez que nada en una tinaja.	Traducción de la descripción en chino: La cámara gira hacia el patio de la familia de <i>Chu</i> .

En este caso concreto no se nombra al señor Chu en el GAD español hasta pasados los primeros 15 minutos de la película, cuando alguien le llama le llama empleando un vocativo.

El siguiente ejemplo también demuestra que las técnicas en ambos GAD difieren a la hora de revelar la identidad de un personaje por primera vez, esta vez el de Jen (Tabla 4).

Tabla 4 Identificación de Jen en su primera aparición

<p>Descripción en español:</p> <p>En un autobús <i>una muchacha</i> golpea su walkman que está fallando.</p>	<p>Descripción en chino: 公交车上，大女儿家珍昂着头在听她的录音机。</p> <hr/> <p>Traducción de la descripción en chino: En el autobús la hija mayor <i>Jen</i> está escuchando a su radio con la cabeza levantada.</p>
---	---

Además, en el GAD chino, se suele nombrar al personaje sin añadir información adicional mientras en el español se incorporan otros detalles, como su relación con otro personaje o una cualidad suya, entre otros. Por ejemplo, en la segunda aparición de Li Kai se dice su nombre directamente en el GAD chino cuando en el español se le relaciona con una característica física mencionada antes (Tabla 5):

Tabla 5 Identificación de Li Kai en su segunda aparición

<p>Descripción en español:</p> <p>El ejecutivo <i>guaperas</i> que dormita.</p>	<p>Descripción en chino: 她转过身，看见了睡在沙发上的李凯。</p> <hr/> <p>Traducción de la descripción en chino: Da media vuelta y ve a Li Kai, quien duerme en el sofá.</p>
--	--

Veamos otros dos ejemplos en el GAD español:

Ejemplo 4: Chu, que *es el padre de las tres hijas*. (se relaciona con otros personajes)

Ejemplo 5: Ming-Dao, *el chico del hombro dislocado*. (se relaciona el personaje con algo ocurrido anteriormente)

Además, hay muchas escenas en las que intervienen varios personajes femeninos a la vez, por ejemplo, cuando cenan juntos las tres hijas y el padre. En la película original las voces de los actores son bastante diferentes, con lo que resultan fácilmente identificables. Sin embargo, en la versión doblada española, las voces se parecen mucho, de modo que podría crear confusión entre el público con pérdida de visión.

Al mismo tiempo, y a diferencia del caso chino, en el GAD español a veces se desvela la identidad del personaje antes de que hable en caso de que vaya a pronunciarse al mismo tiempo que otros. Por ejemplo, esto ocurre en la primera cena dominical, cuando habla Chien con el fin de comunicar la mudanza de la casa vieja en el futuro:

Ejemplo 6: El padre regresó de la cocina con más platos. *Chien*.

4.2. Análisis de los cultuemas

En nuestro estudio, primero hemos etiquetado los cultuemas y los hemos dividido inspirándonos en la clasificación de Molina (2001). Dicha clasificación incluye las siguientes categorías: medio natural (flora, fauna, fenómenos atmosféricos, topónimos, etc.), patrimonio cultural (personajes reales o ficticios, el folklore, obras y monumentos emblemáticos, etc.), cultura social (convenciones y hábitos sociales) y falsos amigos culturales (mismo concepto, comportamiento o gesto con una connotación cultural distinta).

La Tabla 6 presenta la clasificación de los cultuemas encontrados en los dos GADs.

Tabla 6 Clasificación de los cultuemas encontrados

	GAD español	GAD chino
Medio natural	0	1
Patrimonio cultural	72	31
Cultura social	8	2
Falsos amigos culturales	3	0

En nuestro estudio hemos observado que la mayoría de culturemas se aglutinan en la categoría del patrimonio cultural. Esto se puede deber a que los que están incluidos en la categoría “patrimonio cultural” tienden a ser objetos más concretos y variados. Es donde interviene más el audiodescriptor porque generalmente no hay un equivalente léxico en la cultura de llegada. Sin embargo, en la cultura social, como la etiqueta en la mesa, que suele estar formada por una serie de acciones, es más fácil de satisfacer el criterio de objetividad, ya que se suele describir simplemente cuáles son las acciones, sin añadir más comentarios. Finalmente, los falsos amigos son más frecuentes entre lenguas cercanas, por lo que estos son escasos en los GADs estudiados a causa de la lejanía entre las culturas china y la española.

El único elemento bajo la categoría del medio natural es una planta que rodea las paredes de la casa vieja y solo se menciona en el GAD chino: *pashanhu* (爬山虎), ‘parra virgen’, una planta procedente de Asia oriental.

Los culturemas relacionados con la cultura social que hemos detectado giran en torno de, sobre todo, los hábitos gastronómicos. Como en estos casos se trata de acciones llevadas a cabo por algunos personajes, se describe simplemente lo que hacen sin ningún comentario. Por ejemplo, en China existe la costumbre de introducir comida en el cuenco de alguien cuando comen juntos para expresar amabilidad. Esto ocurre con mucha frecuencia en la película analizada. En el GAD español se describe sencillamente:

Ejemplo 7: Ella pone comida con los palillos en el cuenco de Guo Lun.

Un ejemplo del falso amigo es el uso de los boles. Dichos instrumentos se emplean de forma distinta en las dos culturas a pesar de compartir una misma apariencia. En los dos guiones se usa la palabra “bol” en la primera escena para referirse al recipiente donde se guardan unos trozos de panceta (Tabla 7):

Tabla 7 El bol

Descripción en español:	Descripción en chino: 五花肉切片，放在小碗里，放入作料，盖上笼屉，准备开始蒸。
--------------------------------	---

Abre la cesta de bambú donde guardó la comida caliente y escurre el jugo del <i>bol</i> en el aceite hirviendo.	Traducción de la descripción en chino: Corta la panceta frita en trozos. Los pone en un <i>bol</i> pequeño y agrega unos ingredientes. Lo cubre con una tapa en la cesta de bambú y empieza a cocinarlo al vapor.
---	--

En España la palabra “bol” es un préstamo naturalizado de la palabra *bowl* en inglés. Equivale a “cuenco” en castellano. Se usa para beber o para preparar un plato en la cocina, por ejemplo, para hacer una mezcla de ingredientes. Si es de tamaño grande, también puede servir de ensaladera. En cambio, el bol tiene más usos en la gastronomía china dependiendo de su tamaño. Los cuencos de arroz son pequeños y es un utensilio muy importante en la gastronomía china a causa de que el arroz es un alimento básico e imprescindible para la mayoría de los chinos, ya que los cuencos permiten llevarse la comida a la boca con mayor facilidad. Los boles más grandes se usan para servir los platos, sobre todo, los que contienen mucho líquido.

No analizamos aquí los culturemas del patrimonio cultural, porque todos los ejemplos que usaremos en el análisis de las técnicas de traducción para el traslado de los culturemas pertenecen a dicha categoría.

También hemos comprobado, tal y como se muestra en la Tabla 6, que los dos guiones se diferencian en la cantidad de los culturemas descritos: hay muchas más descripciones de culturemas en el GAD español, lo que en nuestra opinión se debe principalmente a la lejanía de la cultura china, que requiere más descripciones de los culturemas en la AD española para que el público español pueda entender aspectos de la otra cultura. Otra diferencia observada es la longitud de las descripciones de los culturemas, que tienden a ser más cortas en el GAD chino. Esto se puede deber, nuevamente, a que son conocimientos que comparten todos los chinos mientras que para los españoles es información nueva que demanda más explicaciones para ser comprendida. Por ejemplo, para la escena donde Chu prepara el almuerzo para Shan, la hija de Jin-Rong (Tabla 8):

Tabla 8 Las costillas de Wuxi

Descripción en español: En casa Chu prepara una salsa que luego vierte sobre unas verduras. Añade unas costillas todo[sic] dentro de una tartera.	Descripción en chino: 老朱家。老朱在给姗姗做便当。放上两块鲜嫩的无锡排骨。
	Traducción de la descripción en chino: En la casa de Chu. Chu prepara el almuerzo. Añade dos costillas de Wuxi, que están tiernas y frescas.

En la versión española, debido al desconocimiento del plato, se describe el proceso de su elaboración señalando los tres ingredientes más importantes: la salsa, unas verduras y unas costillas. En la china, en cambio, lo nombran directamente como “costillas de Wuxi”. Es un plato típico de la gastronomía su (苏), que casi todos los chinos conocen. En este tipo de casos, como se trata de una información compartida, basta con una mención directa y breve, para activar la reconstrucción de la imagen mental concreta.

Una vez detectados los tipos de culturemas en los GADs, procedimos a identificar las técnicas traductoras usadas para su traslado (Tabla 9).

Tabla 9 Técnicas de traducción usadas para el traslado de los culturemas

	GAD español	GAD chino
Adaptación	25	0
Descripción	4	0
Generalización	15	8
Particularización	3	0
Sustitución	1	0

Observamos que las técnicas más usadas en el GAD español son la adaptación y la generalización, mientras que en chino solo hay ocho casos de generalización. A continuación, presentamos dos ejemplos de adaptación y generalización debido a su

frecuente ocurrencia. Asimismo, también mencionamos un caso donde se combinan la adaptación y la particularización.

4.2.1. Adaptación: la técnica culinaria del *baochao* (爆炒)

La adaptación hace referencia a la sustitución de un elemento cultural por otro propio de la cultura receptora. Un ejemplo es el *baochao* (爆炒), una técnica culinaria muy típica de China, que consiste en cocinar a fuego fuerte en un wok. Veamos cómo se ha hecho referencia a esta forma de cocinar en ambos GADs (Tabla 10):

Tabla 10 Baochao (爆炒)

<p>Descripción en español: En casa el padre está <i>flameando</i> verduras y carnes.</p>	<p>Descripción en chino: 老朱家。又一道菜烧好了。 Traducción de la descripción en chino: En la casa de Chu. Ya <i>ha cocinado</i> otro plato.</p>
---	---

Chao (炒) es un método de cocción según el cual se necesita mantener en movimientos constantes los ingredientes en el wok con una temperatura relativamente alta. Se parece mucho a la acción de “saltear” en español. A veces *chao* (炒) se traduce como “saltear en un wok” e incluso incorrectamente como “freír”. En chino el morfema delante de *chao* (炒) especifica aún más cómo se produce este salteado. La palabra *bao* (爆) indica que se precalienta el wok a fuego vivo hasta que se calienta y que luego se le añade una pequeña cantidad de aceite junto con condimentos secos. Cuando emite un olor fragante, se agregan otros ingredientes y se saltean poco tiempo. Generalmente en este paso las llamas envuelven la sartén.

En el GAD chino, se dice simplemente que “está preparado el plato” sin mencionar el proceso de cocción. Se escuchan los sonidos del movimiento de la espátula en el wok, pero es difícil deducir de estos que es *baochao* (爆炒). En el guion español, se traduce como “flamear”, una técnica que utiliza el alcohol sobre la comida para prenderle fuego. También se emplea mucho para la preparación de algunos dulces. La adaptación del método de cocina chino del *baochao* (爆炒) al español flamear se debe a una

característica que comparten en el proceso de la elaboración del plato: el hecho de que se prende fuego a la sartén.

4.2.2. La generalización: el Salón Conmemorativo de Chiang Kai-shek

La generalización se refiere, en términos generales, a la utilización de un término más general o neutro. En cuanto a la AD, como lo que se traducen son elementos visuales, es difícil decidir si se ha usado la técnica de la generalización, ya que en el texto original no se ha definido verbalmente la categoría del componente. En nuestro artículo, definimos esta técnica como la clasificación de un elemento a una categoría superior y más abstracta cuando se ve claramente en la película la categoría inferior a la que pertenece.

Un ejemplo detectado se produce cuando Chu corre por un parque que rodea el Salón Conmemorativo de Chiang Kai-shek, el cual linda con el Teatro Nacional y la Sala de Conciertos Nacional. Ambos son lugares emblemáticos de Taiwán y se distinguen claramente en la escena. En el GAD chino se omite esta información, mientras que en el español se opta por decir:

Ejemplo 8: Después Chu corre por un parque. Hay *bellos templetos* bordeando los paseos.

En realidad, el intervalo temporal disponible es suficiente para profundizar un poco más. Posiblemente la generalización de los dos culturemas como “bellos templetos” se deba a que el autor del GAD español no conoce dichos dos lugares de interés histórico-cultural.

4.2.3. Combinación de la particularización y de la adaptación: el maître

A diferencia de la generalización, definimos la particularización como la clasificación de un elemento en una categoría inferior y más concreta cuando no se capta esta información en la película. En el film, al hablar de la persona que se encarga de los asuntos de la cocina en el restaurante donde ayuda Chu, se usa la palabra *jingli* (经理) ‘gerente’ en el guion chino y “maître” en el español.

En realidad, no se indica en la película el cargo de la persona responsable de la cocina. Tampoco es una información que se pueda obtener a partir de los diálogos entre los personajes. Por lo tanto, consideramos que la interpretación del cargo de la persona referida como *jingli* (经理) responde a la técnica de la particularización en el GAD chino. Por otro lado, mientras que el concepto de “gerente” es compartido por ambas culturas,

el de maître es desconocido por los chinos y constituye un culturema. Es un cargo propio de la cultura occidental que no tiene equivalente en la china. Por consiguiente, creemos que en la traducción del cargo de la persona en cuestión como maître se combinan dos técnicas de traducción: la particularización y la adaptación.

4.2.4. Errores del guion de la AD en español

Cabe mencionar que durante el análisis de los culturemas del GAD español, nos hemos dado cuenta de que, empleando la terminología de Hurtado (2001: 305), hay muchos errores del tipo falso sentido y sin sentido, que afectan a la comprensión del texto original. Según Delisle (1993, citado en Hurtado, 2001: 291):

El falso sentido es “la falta de traducción que resulta de una mala apreciación del sentido de una palabra o de un enunciado en un contexto dado”, sin llegar a causar contrasentido o sin sentido.

El sin sentido es “dar a un segmento del texto de partida una formulación en lengua de llegada totalmente desprovista de sentido o absurda”.

En nuestro caso particular, un falso sentido se produce cuando, por ejemplo, Chu echa *bingtang* (冰糖) ‘azúcar piedra’ sobre un plato. En el GAD español se describe como “cristal de sal gorda”, ya que ambos alimentos se parecen mucho.

En cuanto a los errores de sin sentido, recuperamos el caso de la olla mongola o el hot-pot, uno de los platos más típicos de la gastronomía china. Para prepararlo, se cortan los alimentos y se dejan en unos platos separados y puestos alrededor de la olla, que generalmente está en el centro de la mesa. Cuando el caldo ya está caliente, se escaldan o cuecen dentro los diferentes ingredientes. Una vez cocidos, los alimentos se remojan en la salsa que hay en un cuenco pequeño para comer. A veces se traduce como “fondue china” u “olla mongola”. También se usa para referirse a la olla utilizada. Este plato presenta una serie de variedades en función del tipo de caldo, el equipamiento usado, el método de cocción, el tamaño de la olla y el tipo de la fuente de calor. La que se presenta en la película es el hot-pot tradicional del norte. En chino se denomina *tong huoguo* (铜火锅), donde *tong* (铜) se refiere al material “cobre” de la olla. En la base se coloca el carbón como fuente de calor y en el centro hay una pequeña chimenea por donde sale el humo. No se menciona este plato en el guion chino ni se puede inferir a partir de los diálogos y sonidos de la película. La descripción de esta comida en el guion español es:

Ejemplo 9: Chien cierra el tiro de *un calentador de comida con chimenea*.

Tanto en España como en China un calentador de comida suele hacer referencia a una tartera electrónica, que tiene la función de calentar comida, o al equipamiento usado para mantener la comida caliente en un bufé. Por tanto, las diferencias entre estos dos elementos resultan evidentes: el primero se usa para cocinar y el segundo para calentar los alimentos ya cocinados, pese a que hay un proceso de calentamiento de la comida y el recipiente en ambos casos. Además, la descripción “con chimenea” puede causar aun más confusión entre el público español, porque no es habitual que una tartera electrónica o un calentador de alimentos de un bufé tenga una chimenea. Por otro lado, la olla mongola no solo se refiere a un alimento, sino que también conlleva connotaciones culturales que implican la idea de reunión familiar en la cultura china, puesto que generalmente todos los miembros de la familia se sientan alrededor de la mesa para comer y charlar. En la película, esta comida contrasta con la situación del momento: la salida progresiva de los miembros familiares y, por tanto, el inicio de la disolución familiar. Muy posiblemente los errores de falso sentido o sin sentido sean el resultado del desconocimiento del elemento cultural por parte del audiodescriptor español teniendo en cuenta que la distancia entre la cultura china y la española es bastante grande.

5. Conclusiones

En este artículo hemos comparado, en primer lugar, la forma de describir e identificar los personajes en los GADs analizados. A continuación, hemos analizado el tratamiento y traslado de los culturemas en las ADs de la película escogida y hemos detectado que la mayoría de los culturemas corresponden al patrimonio cultural. Además, se describen más culturemas en el GAD español y las descripciones son más largas. Aparte de esto, las técnicas más usadas para su traslado al español son la adaptación y la generalización. También hemos comentado algunos errores en el GAD español, que pueden ser causados por el desconocimiento de la cultura china por parte del audiodescriptor.

A partir de todo lo mencionado anteriormente, consideramos que hemos verificado las dos hipótesis planteadas al principio de nuestro estudio: la distancia cultural ha desembocado en numerosas disimilitudes en las descripciones de los personajes y de los culturemas. También ha influido en las técnicas para trasladar los culturemas.

Sin embargo, nuestro estudio no está exento de limitaciones. En primer lugar, los resultados no se pueden generalizar debido a que solo se ha estudiado una película y tan

solo dos aspectos de los GADs. Además, como en China el servicio de la AD no es uniforme, las conclusiones que hemos sacado tampoco se pueden aplicar a todas las asociaciones que ofrecen este servicio. En segundo lugar, nos hemos percatado de que la clasificación de técnicas traductoras de Molina y Hurtado (2002) no es la ideal en el caso de la AD porque algunas de ellas no son aplicables para esta modalidad de traducción intersemiótica, como la sustitución (cambio de los elementos lingüísticos por paralingüísticos, o viceversa). Además, Jankowska, Milc y Fryer (2017) también coinciden en que en la actualidad las clasificaciones de los referentes culturales y las técnicas usadas para su traslado en la AD son insuficientes e inconsistentes.

Por lo tanto, creemos que, en el futuro, es imprescindible trabajar con un corpus más grande de películas audiodescritas. Asimismo, sería aconsejable ampliar el análisis y abarcar aspectos tales como las características lingüísticas, las descripciones de las expresiones faciales, el texto en pantalla, el lenguaje cinematográfico, etc. En cuanto al traslado de los culturemas, sería interesante estudiar si las técnicas empleadas actualmente permiten al público ciego captar los culturemas. También sería necesaria una clasificación más completa y precisa de las técnicas para analizar los culturemas.

En cualquier caso, este modesto estudio pretende arrojar un poco de luz acerca de cómo se audiodescribe, sobre todo, en chino y qué aspectos deben tratarse con especial cuidado en caso de que en un futuro se considere seriamente la traducción de GADs del español al chino para poder ampliar la oferta de este servicio de accesibilidad.

Bibliografía citada

ARMA, Saveria, 2016: “Why can’t you wear shoes like the other mothers? Preliminary investigation on the Italian language of audio description” en Elisa PEREGO (ed.): *Emerging topics in translation: audio description*, Trieste: Edizioni Università di Trieste, 37-35.

BARTLETT, Frederic, 1932: *Remembering: A Study in Experimental and Social Psychology*, Cambridge: Cambridge University Press.

BOURNE, Julian y Catalina JIMÉNEZ, 2007: “From the visual to the verbal in two languages: a contrastive analysis of the audio description of *The Hours* in English and

Spanish” en Jorge, DÍAZ-CINTAS, Pilar ORERO y Aline REMAEL (eds.): *Media for All*, Amsterdam: Rodopi, 175-188.

CHAO, Ya-li, 2002: 语言世界中的流动光影。口述影像的理论建构 (Images Flowing in the World of Language: To Construct a Theoretical Framework for Audio Description), Taipei: Wunan Publishing Co.

JANKOWSKA, Anna, 2015: *Translating audio description scripts. Translating as a new strategy of creating audio description*, Frankfurt: Peter Lang.

JANKOWSKA, Anna, Michal MILC y Louise FRYER, 2017: “Translating audio description scripts... into English”, *Skase* 10(2), 2-16.

HURTADO, Amparo, 2011: *Traducción y traductología: introducción a la traductología*, Madrid: Cátedra.

LEUNG, Hoi Ching Dawning, 2018: *Audio description of audiovisual programmes for the visually impaired in Hong Kong*. Tesis doctoral, University College London.

LI, Dongxiao (李东晓), 2013, 听见, 看见: 影视媒体的无障碍传播研究 (Hear and see: audiovisual media for all), Hangzhou: Zhejiang University Press.

LIMBACH, Christiane, 2012: *La audiodescripción fílmica y el aspecto de la neutralidad. Un análisis comparativo entre la audiodescripción en inglés y alemán de la película “Slumdog millionaire”*. Tesis doctoral, Universidad de Granada.

LÓPEZ VERA, Juan Francisco, 2006: “Translating audio description scripts. The way forward? Tentative first stage project results” en Mary CARROLL, Heidrun GERZYMISCH-ARBOGAST y Sandra NAUERT (eds.): *MuTra 2006 – Audiovisual Translation Scenarios Conference Proceedings*: Saarbrücken, 1-10.

MATAMALA, Anna y Naila RAMI, 2009: “Análisis comparativo de la audiodescripción española y alemana de ‘Good-bye Lenin’”, *Hermeneus* 11, 249-266.

MATAMALA, Anna, 2018: “One short film, different audio descriptions. Analysing the language of audio descriptions created by students and professionals”, *Onomázein* 41, 185-207.

MAYORAL, Roberto, 1994: “La explicitación de información en la traducción intercultural” en Amparo HURTADO (ed.): *Estudis sobre la traducció*, Castelló de la Plana: Universitat Jaume I, 73-96.

MOLINA, Lucía y Amparo HURTADO, 2002: “Translation techniques revisited: a dynamic and functionalist approach” en *Meta*: 47/4, 398-512.

MOLINA, Lucía, 2001: *Análisis descriptivo de la traducción de los culturemas árabe-español*. Tesis doctoral, Universitat Autònoma de Barcelona.

NEWMARK, Peter, 1988: *A textbook of Translation*, London: Prentice-Hall.

NORD, Christiane, 1994: “It’s tea-time in Wonderland. Cultura-makers in fictional texts” en *Intercultural Communication: proceedings of the 17th International LAUD Symposium*, Duisburg.

NORD, Christiane, 1997: *Translating as a Purposeful Activity: Functionalist Approaches Explained*, Manchester: St. Jerome Publishing.

SANTAMARÍA, Laura, 2001: “Culture and translation. The referential and expressive value of cultural references” en Rosa AGOST y Frederic CHAUME (eds.): *La traducción en los medios audiovisuales*, Castelló de la Plana: Universitat Jaume I, 159-164.

TOR-CARROGGIO, Irene y Helena CASAS-TOST (próximamente): “Who Is Currently Audio Describing in China? A Study on the Profile of Chinese Audio Describers”. Aparecerá en *MonTI. Monografías de Traducción e Interpretación* 12.

TOR-CARROGGIO, Irene, próximamente: “T(ime) T(o) S(tart) Synthesising AD in China? Results of a Reception Study”.

TOR-CARROGGIO, Irene y Gert VERCAUTEREN (próximamente): “When East Meets West: A Comparison of Audio Description Guidelines in China and Europe”.

SERGEI, Vlahov y Florin SIDER, 1970: “Neperevodimoe v perevode: realii” en *Masterstvo perevoda 1969*, Moscú: Sovetskii pisatel, 432-456.

WU Zongyi (吴宗艺) y Zhenzhen XIE (谢桢桢), 2015: «Zhongguo Shizhang Koushuyingxiang fuwu de fazhan xianzhuang yu dalu tuiguang (中国视障口述影像服务的发展现状与大陆推广)» [La situación actual del servicio de audiodescripción de

películas en China y su promoción en la China continental], *Xinwen Yanjiu Daokan* (新闻研究导刊) [Revista de investigaciones periodísticas], 6 (10), 21.

YEUNG, Jessica, 2007: "Audio description in the Chinese world" en Jorge DÍAZ-CINTAS, Pilar ORERO y Aline REMAEL (eds.): *Media for All*, Amsterdam: Rodopi, 231-244.

Annex 3: Ethics Committee Documentation

3.1. Procedure approved for the study with audio describers

Informació requerida per la CEEAH de la UAB

Num. CEEAH: **4133**

Data: **20-03-2018**

1. Títol del procediment de recerca

El perfil dels audiodescriptors a la Xina

2. Breu descripció del projecte

El compromís de la societat envers l'accessibilitat està enfortint-se arreu del món, sobretot arran de l'aprovació de la Convenció sobre els Drets de Persones amb Discapacitats per part de les Nacions Unides el 2008, que emfasitza especialment el dret a la informació i a la comunicació, àrees en les quals la traducció audiovisual té una importància clau. En els darrers anys, els continguts audiovisuals són potser un dels àmbits en els quals es fan més esforços en aquest sentit, ja que estan experimentant una gran expansió i han de ser adaptats a persones amb tot tipus de discapacitats sensorials als teatres, cinemes, museus, televisions, etc. Gràcies a serveis com ara la subtitulació per a sords i l'audiodescripció per a persones oegues o amb baixa visió, aquestes persones poden accedir a informació que, d'una altra manera, els seria impossible de gaudir.

Cal destacar que, tot i que l'audiodescripció aplicada a productes audiovisuals ja compta amb uns quants anys de trajectòria en l'àmbit professional i en de recerca a Europa i als EUA, a la Xina és, en comparació, un àmbit que encara es troba en estat embrionari, malgrat ser el país amb més persones discapacitades del món segons l'Organització Mundial de la Salut. Val a dir que l'oferta d'aquest servei d'accessibilitat al país és irregular, és a dir, com que no existeix una legislació concreta sobre aquest tema, cada ciutat l'ofereix o no segons les seves possibilitats (cal recordar que les diferències socioeconòmiques entre ciutats a la Xina poden ser molt grans). No obstant això, cal aclarir que sí que existeix una llei sobre accessibilitat (Llei sobre Protecció de les Persones Discapacitades de la República Popular de la Xina, aprovada el 1990 i revisada el 2008), l'article 43 del Capítol 5 de la qual resulta especialment pertinent de cara a la tesi doctoral en què s'inscriu aquest projecte, atès que justifica la recerca que proposem: "El govern i la societat han d'adoptar les següents mesures per enriquir la vida espiritual i cultural de les persones discapacitades: (#) Oferir programes de televisió en llengua de signes, programes de ràdio especialment dissenyats per a persones discapacitades i afegir subtítols o narracions a més programes de televisió i pel·lícules".

Tot i que l'audiodescripció existeix a la Xina des de fa uns anys i ha aconseguit fites gens menyspreables pel que fa a l'abast del servei a ciutats com ara Shanghai i Hong Kong, encara no s'ha abordat en l'àmbit acadèmic de forma àmplia.

Aquest projecte s'emmarca dins de la tesi doctoral de la doctoranda Irene Tor Carroggio, un dels objectius de la qual és conèixer l'estat de la qüestió de l'audiodescripció a la Xina i, més concretament, quin és el perfil dels audiodescriptors del país.

3. Dades de l'investigador responsable

Nom i cognoms	Helena Casas Tost
NIF	38724891N
Departament / Centre	Dept. Traducció
Telèfon	(+34)935813372
Adreça electrònica	Helena.Casas@uab.cat
Acreditació	No

4. Objectius del procediment d'experimentació amb humans

Descriure els principals objectius que es pretenen assolir amb la realització d'aquest procediment d'experimentació

- a) Identificar el perfil dels audiodescriptors de Shanghai, la ciutat pionera en audiodescripció a la Xina;
- b) comparar-lo amb el d'altres audiodescriptors provinents de diferents ciutats de la Xina;
- c) i comparar-lo amb el perfil dels audiodescriptors dels països europeus que van respondre el qüestionari del projecte europeu ADLABPRO (<https://adlabpro.wordpress.com/>).

5. Metodologia del procediment d'experimentació

Descriure breument la metodologia emprada justificant les dades, mostres biològiques i o respostes conductuals obtingudes de les persones sota experimentació

L'anàlisi del perfil dels audiodescriptors es farà mitjançant una única via de recollida de dades: un qüestionari en línia basat en l'emprat per al mateix objectiu pertanyent al projecte ADLABPRO. A més de redactar-lo en xinès, s'hi han incorporat lleugeres modificacions, aprovades pels líders del projecte ja esmentat (Dra. Elisa Perego, Universitat de Trieste) per tal d'adaptar-se a la realitat que de moment es coneix de la Xina i per assegurar-nos que siguin comparables. El qüestionari ha estat creat amb Web Surey Creator, perquè és accessible des de la Xina. L'enllaç al qüestionari és: <http://survey.websurveycreator.com/s.aspx?s=35b02c8b-6d5b-45d0-87a9-5d9b5ae19e2d>

A l'adjunt hi trobareu el qüestionari en versió bilingüe anglès/xinès.

[Veure Annex 3 \(application/pdf - 607.666015625 kB\)](#)

6. Informació a les persones participants

Supòsit	Si	No	
S'annexa un full d'informació del projecte de recerca que inclou de forma entenedora els objectius de la investigació, els investigadors/res responsables i la forma d'obtenir fàcilment més informació?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Adjunt
S'annexa un full de consentiment informat signat per l'investigador/a i la persona en qüestió on queda clarament expressat que la participació és voluntària, que disposa de la informació suficient i que es podrà retirar en qualsevol moment sense donar explicacions i sense que això tingui cap mena de conseqüència?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Adjunt

7. Compensació

Supòsit	Si	No
Està previst algun tipus de compensació per la participació en el projecte?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, explicar i justificar-ne les característiques, la quantia (si és econòmic) i l'adequació amb el risc i/o molèstia ocasionats al subjecte participant</i>		

8. Gestió i emmagatzematge de les dades obtingudes

Supòsit	Si	No
Està prevista l'anonimització de les dades obtingudes? Recordeu que s'entén per anonimització la desvinculació permanent i irreversible entre les dades i la identitat del subjecte de recerca	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<i>Si la resposta es Si, explicar les activitats realitzades</i>		
Les dades són anònimes, atès que els qüestionaris no tenen cap dada que permeti identificar els informants.		
Està previst l'emmagatzematge de les dades amb mesures de seguretat?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, donar detalls del procediment de seguretat</i>		

9. Feedback

Supòsit	Si	No
Està prevista alguna forma de feedback a les persones participant un cop finalitzat el projecte?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, quin?</i>		
No es farà arribar feedback directament als informants. Amb tot, es farà arribar l'article amb els resultats obtinguts a les associacions i persones que hagin participat i que ens hagin ajudat a localitzar els audiodescriptors i a		

distribuir els qüestionaris en línia, però no tenim les dades de contacte dels informants, per tant no ens hi podem adreçar personalment. Al qüestionari oferim el correu de la recercadora principal per tal que, en cas de voler cap informació, s'hi puguin adreçar directament.

10. Registre dades

Supòsit	Si	No
Les dades obtingudes es troben en un registre aprovat per l'Autoritat Catalana de Protecció de Dades?	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Si la resposta es Si, indicar codi del registre:

Nom i cognoms

Lloc i data

Signatura

3.2. Approval for the study with audio describers



Universitat Autònoma de Barcelona

Vicerectorat d'Investigació

Comisión de Ética en la Experimentación Animal y Humana (CEEAH)

Universitat Autònoma de Barcelona
08193 Bellaterra (Cerdanyola del Vallès)

La Comisión de Ética en la Experimentación Animal y Humana (CEEAH) de la Universitat Autònoma de Barcelona, reunida el día **23-03-2018**, acuerda informar favorablemente el proyecto titulado "**El perfil dels audiodescriptors a la Xina**" presentado por **Helena Casas Tost**

Elaborado:	Aprobado:
<p>Nombre: Nuria Perez Pastor Cargo: Secretària de la CEEA de la UAB Fecha:</p> <p>NURIA PEREZ PASTOR</p> <p><small>Firmado digitalmente por NURIA PEREZ PASTOR Nombre de reconocimiento (DN): c=ES, ou=Regio https://www.uaoz.cat/, CAT/Cat@Regulacio, sn=PEREZ PASTOR, givenName=NURIA, serialNumber=351096387, cn=NURIA PEREZ PASTOR Fecha: 2018.04.09 08:37:42 +02'00'</small></p>	<p>Nombre: José Luis Molina González Cargo: President de la CEEAH de la UAB Fecha:</p> <p>MOLINA GONZALEZ, JOSE LUIS (AUTENTICACIÓN)</p> <p><small>Firmado digitalmente por MOLINA GONZALEZ, JOSE LUIS (AUTENTICACIÓN) Fecha: 2018.04.06 09:44:30 +02'00'</small></p>

3.3. Procedure approved for the study with users

Informació requerida per la CEEAH de la UAB

Num. CEEAH: 4258

Data: 26-04-2018

1. Títol del procediment de recerca

Opinió dels usuaris cecs o amb baixa visió xinesos sobre l'audiodescripció oferta actualment i valoració de l'ús de veus sintètiques per poder augmentar-ne l'oferta

2. Breu descripció del projecte

El compromís de la societat envers l'accessibilitat està enfortint-se arreu del món, sobretot arran de l'aprovació de la Convenció sobre els Drets de Persones amb Discapacitats per part de les Nacions Unides el 2006, que emfasitza especialment el dret a la informació i a la comunicació, àrees en les quals la traducció audiovisual té una importància clau. En els darrers anys, els continguts audiovisuals són potser un dels àmbits en els quals es fan més esforços en aquest sentit, ja que estan experimentant una gran expansió i han de ser adaptats a persones amb tot tipus de discapacitats sensorials als teatres, cinemes, museus, televisions, etc. Gràcies a serveis com ara la subtítolació per a sords i l'audiodescripció per a persones cegues o amb baixa visió, aquestes persones poden accedir a informació que, d'una altra manera, els seria impossible de gaudir.[]

A diferència de països com ara Espanya, el Regne Unit i els Estats Units, l'audiodescripció aplicada a productes audiovisuals encara es troba en estat embrionari a la Xina, malgrat ser el país amb més persones discapacitades del món segons l'Organització Mundial de la Salut. L'oferta d'aquest servei d'accessibilitat al país és irregular, és a dir, tot i que existeix la Llei sobre Protecció de les Persones Discapacitades de la República Popular de la Xina ꟷla qual encoratja el govern i la societat a "oferir programes de televisió en llengua de signes, programes de ràdio especialment dissenyats per a persones discapacitades i afegir subtítols o narracions a més programes de televisió i pel·lícules"ꟷ, cada ciutat l'ofereix o no en funció de les seves possibilitats, que varien enormement arreu del país.

Tot i que l'audiodescripció existeix a la Xina des de fa uns anys i ha aconseguit fites gens menyspreables pel que fa a l'abast del servei a ciutats com ara Shanghai i Hong Kong, encara no s'ha abordat en l'àmbit acadèmic de forma àmplia.

Aquest projecte s'emmarca dins de la tesi doctoral de la doctoranda Irene Tor Carroggio, un dels objectius de la qual és conèixer els hàbits de consum i la satisfacció dels usuaris xinesos cecs o amb baixa visió pel que fa a l'audiodescripció a què tenen accés. A més, s'aprofitarà que encara ens trobem en un estadi inicial del projecte per preguntar per l'interès que els desperta la proposta que pretén investigar la tesi en fases posteriors: esbrinar si és possible l'acceptació de veus sintètiques aplicades a l'audiodescripció en xinès.

Atès que la tesi adopta un enfocament centrat en els usuaris, la seva participació és imprescindible al llarg de diferents estadis de la tesi, des de l'inici fins al final. En aquesta fase és important per tal que mostrin si la línia de recerca adoptada a la tesi els interessa i troben que pot tenir un impacte social real i perquè aportin dades per tal de definir els experiments subsegüents, que tenen per objectiu satisfer les seves necessitats com a usuaris final i reals de l'audiodescripció.

Àrea del procediment d'experimentació amb humans

Humanitats

3. Dades de l'investigador responsable

Nom i cognoms	Helena Casas Tost
NIF	39724991N
Departament / Centre	Dept. Traducció

Telèfon	(+34)935813372
Adreça electrònica	Helena.Casas@uab.cat
Accreditació	No

4. Objectius del procediment d'experimentació amb humans

Descriure els principals objectius que es pretenen assolir amb la realització d'aquest procediment d'experimentació

Conèixer els hàbits de consum, l'experiència amb l'audiodescripció, així com l'opinió dels usuaris cecs i amb baixa visió xinesos sobre noves opcions per augmentar l'oferta del servei, en concret, sobre la possibilitat de fer servir veus sintètiques per audiodescriure pel·lícules.

5. Metodologia del procediment d'experimentació

Descriure breument la metodologia emprada justificant les dades, mostres biològiques i o respostes conductuals obtingudes de les persones sota experimentació

L'anàlisi de la satisfacció i del consum d'audiodescripció dels usuaris es farà mitjançant dues vies de recollida de dades, que corresponen a dues fases diferents en el temps: d'una banda, a través d'un qüestionari distribuït en persona a Shanghai i en línia (per tal de poder arribar a usuaris d'altres llocs de la Xina); i, de l'altra, a través de dos grups de discussió l'objectiu dels quals serà aprofundir en els resultats més destacats dels qüestionaris o en els punts que calgui aclarir.

La metodologia s'ha plantejat d'aquesta manera perquè, per a altres estudis realitzats a la Xina amb cecs, sabem que és molt possible que hi hagi algunes preguntes que no acabin d'entendre, malgrat que ens hem esforçat a redactar els qüestionaris de la forma més entenedora possible. Cal recordar que és molt possible que els informants siguin persones grans o sense gaire formació, per la qual cosa com més ocasions tinguem per preguntar-los, més probabilitats tenim de poder-nos entendre i d'obtenir feedback rellevant.

A l'adjunt hi trobareu el qüestionari en versió bilingüe castellà/xinès, en el qual s'integra també el full de consentiment informat (com que només ens deixa adjunta un arxiu, aquí hi posem la versió en xinès i en el punt 8 la traducció al castellà). La distribució dels qüestionaris a Shanghai es farà en persona al cinema Cathay, on l'últim dijous de cada mes se celebra una sessió de cinema audiodescrit. Cinc voluntaris xinesos col·laboraran llegint el contingut del qüestionari als usuaris per tal que puguin contestar oralment.

El contingut concret dels grups de discussió no es pot definir encara perquè depèn del que trobem a l'estadi anterior. Tot i així, la idea de les recercadores és aprofundir en els aspectes més qualitius del qüestionari com ara què valoren i què troben a faltar a les audiodescripcions existents, en quins altres àmbits estarien interessats a aplicar aquest servei i recollir informació o la seva opinió sobre les noves àrees que voldrien que es desenvolupessin més.

[Veure Annex 3 \(application/pdf - 134.9091796875 kB\)](#)

6. Informació a les persones participants

Supòsit	Si	No	
S'annexa un full d'informació del projecte de recerca que inclou de forma entenedora els objectius de la investigació, els investigadors/res responsables i la forma d'obtenir fàcilment més informació?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Annex 1
S'annexa un full de consentiment informat signat per l'investigador/a i la persona en qüestió on queda clarament expressat que la participació és voluntària, que disposa de la informació suficient i que es podrà retirar en qualsevol moment sense donar explicacions i sense que això tingui cap mena de conseqüència?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Annex 2

7. Compensació

Supòsit	Si	No
---------	----	----

Està previst algun tipus de compensació per la participació en el projecte?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, explicar i justificar-ne les característiques, la quantia (si és econòmic) i l'adequació amb el risc i/o molèstia ocasionats al subjecte participant</i>		

8. Gestió i emmagatzematge de les dades obtingudes

Supòsit	Si	No
Està prevista l'anonimització de les dades obtingudes? Recordeu que s'entén per anonimització la desvinculació permanent i irreversible entre les dades i la identitat del subjecte de recerca	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<i>Si la resposta es Si, explicar les activitats realitzades</i>		
Les dades són anònimes, atès que els qüestionaris no tenen cap dada que permeti identificar els informants		
Està previst l'emmagatzematge de les dades amb mesures de seguretat?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, donar detalls del procediment de seguretat</i>		

9. Feedback

Supòsit	Si	No
Està prevista alguna forma de feedback a les persones participant un cop finalitzat el projecte?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, quin?</i>		
No es farà arribar feedback directament als informants. No obstant això, es farà arribar l'article amb els resultats obtinguts a les associacions i persones que hagin participat i que ens hagin ajudat a localitzar els usuaris i a distribuir els qüestionaris en línia i en persona, però no tenim les dades de contacte dels informants, per tant no ens hi podem adreçar personalment. Al qüestionari oferim la possibilitat de posar-se en contacte amb la doctoranda i les seves dues directores per més informació o per demanar els resultats un cop analitzades les dades.		

10. Registre dades

Supòsit	Si	No
Les dades obtingudes es troben en un registre aprovat per l'Autoritat Catalana de Protecció de Dades?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, indicar codi del registre:</i>		

Nom i cognoms

Lloc i data

Signatura

3.4. Approval for the study with users



Universitat Autònoma de Barcelona

Vicerectorat d'Investigació

Comisión de Ética en la Experimentación Animal y Humana (CEEAH)

Universitat Autònoma de Barcelona
08193 Bellaterra (Cerdanyola del Vallès)

La Comisión de Ética en la Experimentación Animal y Humana (CEEAH) de la Universitat Autònoma de Barcelona, reunida el día **22-06-2018**, acuerda informar favorablemente el proyecto titulado "**Opinió dels usuaris cecs o amb baixa visió xinesos sobre l'audiodescripció oferta actualment i valoració de l'ús de veus sintètiques per poder augmentar-ne l'oferta**" presentado por **Helena Casas Tost**

<p>Elaborado:</p> <p>Nombre: Nuria Perez Pastor Cargo: Secretària de la CEEA de la UAB Fecha:</p> <p>NURIA PEREZ PASTOR</p> <p><small>Firmado digitalmente por NURIA PEREZ PASTOR Nombre de reconocimiento (DN): cn=ES, ou=Vigne/https://www.dic.cat/CAT/Cert/Regulacio, sn=Perez PASTOR, givenName=NURIA, serialNumber=35109638E, cn=NURIA PEREZ PASTOR Fecha: 2018.07.04 11:25:11 +02'00'</small></p>	<p>Aprobado:</p> <p>Nombre: José Luis Molina González Cargo: President de la CEEAH de la UAB Fecha:</p> <p>MOLINA GONZALEZ, JOSE LUIS (FIRMA)</p> <p><small>Firmado digitalmente por MOLINA GONZALEZ, JOSE LUIS (FIRMA) Fecha: 2018.06.29 16:11:27 +02'00'</small></p>
---	--

3.5. Procedure approved for the reception study

Informació requerida per la CEEAH de la UAB

Num. CEEAH: 4759

Data: 31-05-2019

1. Títol del procediment de recerca

Estudi de recepció sobre l'ús de veus sintètiques en l'audiodescripció (AD) en xinès

2. Breu descripció del projecte

Aquest projecte s'emmarca dins de la tesi doctoral d'Irene Tor Carroggio. Aquesta tesi doctoral estudia, en primer lloc, l'estat de la qüestió de l'AD de pel·lícules a la Xina des d'angles diferents: dibuixa el perfil dels audiodescriptors, recull les necessitats i el grau de satisfacció dels usuaris, i també presenta i analitza algunes de les guies de bones pràctiques amb les quals s'elabora aquest servei d'accessibilitat. La segona part de la tesi consisteix a dur a terme un estudi de recepció amb usuaris xinesos per estudiar l'acceptació de les veus sintètiques en xinès aplicades a l'AD. En cas que el resultat sigui favorable, s'obririen noves possibilitats per a aquest servei d'accessibilitat al país, ja que l'ús d'aquest tipus de tecnologies abarateixen els costos i n'augmenten la velocitat de producció.

3. Dades de l'investigador responsable

Nom i cognoms	Sara Rovira Esteva
NIF	46728204R
Departament / Centre	Dept. Traducció
Telèfon	(+34)935813389
Adreça electrònica	Sara.Rovira@uab.cat
És doctor/a? (recordeu que la comissió només avalua projecte de recerca dirigits per doctors/es)	Si

4. Objectius del procediment d'experimentació amb humans

Descriure els principals objectius que es pretenen assolir amb la realització d'aquest procediment d'experimentació

- Avaluar i comparar la qualitat de dues veus (una de sintètica i una altra d'humana) aplicades en l'AD en xinès d'una pel·lícula històrica.
- Conèixer si les persones cegues xineses acceptarien les veus sintètiques quan escolten una AD d'un producte audiovisual com a solució temporal per tal de tenir més materials d'aquest tipus audiodescrits.
- Conèixer si les persones cegues xineses acceptarien les veus sintètiques quan escolten una AD d'un producte audiovisual de forma permanent.

5. Metodologia del procediment d'experimentació

Descriure breument la metodologia emprada justificant les dades, mostres biològiques i o respostes conductuals obtingudes de les persones sota experimentació

a. Previ al procediment:

El procediment es farà en persona. Abans de començar, es llegirà el full d'informació i es demanarà que els usuaris donin el seu consentiment informat (vegeu apèndix 1) oralment. El consentiment s'enregistrarà en format d'àudio.

Durant el procediment:

S'han triat dos clips (A i B) de gairebé 5 minuts d'una mateixa pel·lícula xinesa (Our time will come, dirigida per Ann Hui). El tipus de pel·lícula (històrica) amb què es fa l'experiment va ser triat pels usuaris en un procediment anterior (aprovat per aquesta comissió) per tal de fer la recerca més centrada en ells. Cada usuari mirarà aquests dos clips: el primer estarà audiodescrit amb una veu sintètica (masculina), mentre que el segon estarà audiodescrit amb una veu humana (també masculina). L'ordre dels clips serà aleatori, però es presentarà sempre primer la veu sintètica.

Després d'escoltar cada clip, els usuaris hauran de contestar un qüestionari que inclou tres preguntes de comprensió, així com una sèrie de preguntes per valorar alguns paràmetres de la veu escoltada que serveixen per avaluar la seva qualitat (vegeu apèndix 2). Per tant, cada usuari contestarà dos qüestionaris, un per a cada clip. Al final del procediment, els usuaris hauran de contestar un últim qüestionari amb preguntes demogràfiques i de preferències (vegeu apèndix 3).

Els tres qüestionaris es llegiran en veu alta i la recercadora anirà prenent nota de les respostes dels usuaris.

Les persones es reclutaran mitjançant contactes personals de la recercadora i l'associació de persones discapacitades de Shanghai. S'esperen aconseguir, com a mínim, 30 usuaris. Els visionats es duran a terme en una sala de reunions de la Universitat Huadong Shifan Daxue (Shanghai) i també als domicilis d'alguns usuaris.

Atès que ni la universitat ni les associacions de cecs col·laboradores no disposen de cap comitè d'ètica no es pot aportar cap aval institucional en aquest sentit. Això no obstant, la investigadora aportarà a les properes setmanes una carta de recolzaments al projecte i al disseny experimental de la seva tutora a la universitat d'acollida.

6. Informació a les persones participants

Supòsit	Si	No	
S'annexa un full d'informació del projecte de recerca que inclou de forma entenedora els objectius de la investigació, els investigadors/res responsables i la forma d'obtenir fàcilment més informació?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Adjunt
S'annexa un full de consentiment informat signat per l'investigador/a i la persona en qüestió on queda clarament expressat que la participació és voluntària, que disposa de la informació suficient i que es podrà retirar en qualsevol moment sense donar explicacions i sense que això tingui cap mena de conseqüència?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Adjunt

7. Compensació

Supòsit	Si	No
Està previst algun tipus de compensació per la participació en el projecte?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<i>Si la resposta es Si, explicar i justificar-ne les característiques, la quantia (si és econòmic) i l'adequació amb el risc i/o molèstia ocasionats al subjecte participant</i>		

8. Gestió i emmagatzematge de les dades obtingudes

Supòsit	Si	No
Està prevista l'anonimització de les dades obtingudes? Recordeu que s'entén per anonimització la desvinculació permanent i irreversible entre les dades i la identitat del subjecte de recerca	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<i>Si la resposta es Si, explicar les activitats realitzades</i> Les dades ja es recolliran de manera anònima, per tant no caldrà que la investigadora les anonimitzi.		
El qüestionari amb dades personals es guardarà de forma separada i es destruirà al finalitzar la tesi.		
Està previst l'emmagatzematge de les dades amb mesures de seguretat?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<i>Si la resposta es Si, donar detalls del procediment de seguretat</i> Les dades dels usuaris, per les quals es pregunta al final del procediment, s'emmagatzemaran en forma de base de dades en un espai virtual d'emmagatzematge de la UAB (Nebula). Els qüestionaris en paper que la recercadora emplei a mesura que els usuaris contestin oralment es guardaran al despatx 128 de l'edifici MRA de la UAB, que roman tancat amb clau. Els arxius d'àudio que es deriven de la recerca es guardaran també al Nebula.		
Un cop defensada la tesi tots aquests materials es destruiran definitivament.		

9. Feedback

Supòsit	Si	No
Està prevista alguna forma de feedback a les persones participant un cop finalitzat el projecte? <i>Si la resposta es Si, quina?</i>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Es redactarà una nota de premsa per a les associacions implicades per tal que la puguin enviar als seus usuaris.		

10. Registre dades

Supòsit	Si	No
Les dades obtingudes es troben en un registre aprovat per l'Autoritat Catalana de Protecció de Dades? <i>Si la resposta es Si, indicar codi del registre:</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Nom i cognoms Lloc i data

Signatura

3.6. Approval for the reception study

**Comisión de Ética en la Experimentación Animal y Humana (CEEAH)**

Universitat Autònoma de Barcelona
08193 Bellaterra (Cerdanyola del Vallès)

La Comisión de Ética en la Experimentación Animal y Humana (CEEAH) de la Universitat Autònoma de Barcelona, reunida el día **28-06-2019**, acuerda informar favorablemente el proyecto con número de referencia **CEEAH 4759** y que tiene por título "**Estudi de recepció sobre l'ús de veus sintètiques en l'audiodescripció (AD) en xinès**" presentado por **Sara Rovira Esteva**

<p>Elaborado:</p> <p>Nombre: Nuria Perez Pastor Cargo: Secretària de la CEEA de la UAB Fecha: Firmado digitalmente por NURIA PEREZ PASTOR Fecha: 2019.07.05 13:49:24 +02'00'</p>	<p>Aprovado:</p> <p>Nombre: José Luis Molina González Cargo: President de la CEEAH de la UAB Fecha: JOSE LUIS MOLINA GONZALEZ - DNI 36561625C Firmado digitalmente por JOSE LUIS MOLINA GONZALEZ - DNI 36561625C Fecha: 2019.07.05 17:43:21 +02'00'</p>
---	--

3.7. Informed consent form for the study with audio describers (Chinese)

该调查问卷是欧洲项目活动 ADLAB PRO 的一部分，该项目旨在为专业音频描述者（编剧以及解说员）的工作提供培训材料。您为此付出的时间和作出的贡献对我们的研究具有重要和宝贵的价值，在此我们十分感谢您的参与。

参与条款

完成本问卷表示您同意接受我们的调查，而更重要的是，您已知晓我们将对您的作答严格保密，并会对您的回答进行编码，以便在将来的出版物和演示文稿中保持匿名。同时，您有权随时放弃此次调查而无需接受任何处罚。

说明

完成此问卷大约需要 20 分钟。问卷分为五个部分，每一部分我们都预留了额外的空间，能够让您对您认为重要而我们却未讨论到的内容进行评论。

3.8. Informed consent form for the study with audio describers (English)

This questionnaire is part of the working activities of the European project ADLAB PRO, which is designed to produce training material for the creation of the professional audio describer. Your time and contribution to our research are invaluable, and we would like to thank you for accepting to participate as a respondent.

TERMS OF PARTICIPATION

Please note that completing the questionnaire implies consent. It is important that you know that all of your responses will be strictly confidential and will be encoded in order to keep your anonymity in future publications and presentations. You have the right to abandon the survey at any time without any penalty.

INSTRUCTIONS

Completing the questionnaire should take approximately 20 minutes. The questionnaire is divided into sections. All sections have additional space for you to include comments on anything you consider of importance and which we have not addressed.

3.9. Informed consent form for the study with users (Chinese)

研究情况介绍

口述影像是一种通过语言描述使视力障碍者最大程度获取特定影像信息的方法。该调查问卷构成博士论文的一部分，主要研究合成语音（例如屏幕阅读器和百度地图等软件中的声音）在电影口述影像中的应用。该服务在上海也被称之为“无

障碍电影”。本问卷旨在调查用户对口述影像服务的体验，了解其对不同改进措施的看法，以便进行不断提高。

本问卷仅做学术研究之用。要求参与对象必需是来自中国的电影口述影像使用者，视力全盲或局部失明。完成此问卷大约需要 10 分钟，调查者应尽力以真实情况作答。

该调查不含任何物质奖励，参与者均采取自愿参与的方式。您有权随时退出此次活动，且不会对您产生任何负面影响。您所填写的内容将被匿名处理。如果您已清楚了解参与条件并同意接受调查，欢迎进入接下来的答题环节。该项目的负责人能将您所填写的内容分享跟其他研究者用于研究。

如果您有任何疑问，请致电联系研究员董琳娜，电话联系方式为 13162099593，或选择向以下邮箱 irene.tor@uab.cat/ sara.rovira@uab.cat/helena.casas@uab.cat 发送邮件。

日期:

负责人签字:

帮助者签字:

您是否完全了解并口头同意参与本次调查？如果您的答复是肯定的，您就说“是”，然后我们就会开始录下您的回答。

是

否

3.10. Informed consent form for the study with users (English)

This study belongs to a PhD research that aims at investigating the possibility of offering audio description using synthetic voices instead of a human voice. Advances in technology mean synthetic voices are improving quickly and the reason we are exploring this option is that this could be a way to produce more audio described films more quickly.

However, the aim is not to replace human voices in audio description but to find out whether Chinese AD users accept synthetic voices in film audio description.

The test, which will last approximately 40 minutes, consists of two parts. First you will listen to one clip that lasts around 5 minutes. Then you will have to answer some questions about it. After that, you will listen to a second clip of a similar duration and will have to answer some more questions. The two clips chosen come from Ann Hui's historical film "Our time will come". This movie revolves around the resistance movement during Japan's occupation of Hong Kong.

The researcher in charge of the study is Mrs. Irene Tor. You can contact her at 13162099635 and ask her for more information about the study and the results.

Now please listen to the consent form:

Your participation in the tests is absolutely voluntary, and there is no economic compensation. The information you provide will be used in the study, but it will remain anonymous. All information that you give will be treated in the strictest confidence and it will only be used for the purposes of this doctoral research. No personal data will be collected that permits your identification. This audio consent will be kept in a safe place by the main researcher and according to the European law will be destroyed after five years once the investigation has finished. You can discontinue your involvement in the study at any time without prior justification. This shall have no repercussions or negative consequences of any sort for you.

If you are willing to participate, please confirm the following statements by saying "yes" at the end. If you say "no" it means you do not give your consent.

- I have been read and have understood the information given for this research;
- I have had the opportunity to ask questions about this research;
- I consent to take part in the research sessions.

3.11. Informed consent form for the reception study (Chinese)

欢迎欢迎参加这次试验。

这项研究是一篇博士论文的一部分，它的目的是研究我们是否有可能在无障碍电影中用机器人语音来替代人声。

科技的进步使得机器人语音也在迅速完善。我们之所以选择探讨这个问题，是因为如果能用机器人语音来替代人声，那么我们能更快、更多地制作无障碍电影。但是，我们研究的目的并不是真的要实现这种替代，而是想探讨中国观看无障碍电影的人群能否接受无障碍电影中出现机器人语音。

这次试验一共有 2 部分，总共会持续大约 40 分钟。首先您会听到一段 5 分钟左右的片段，之后您需要根据听到的片段回答一些问题。接下来，您会听到另一段 5 分钟左右的片段，同样地，您也要回答一些问题。这两个片段都选自香港导演许鞍华的电影《明月几时有》，讲的是日本占领香港期间人民群众发起抵抗运动的故事。

这次研究的负责人是董琳娜女士。您如果想更深入地了解这项研究，或是对试验结果感兴趣，可以随时联系她，她的电话是 13162099635。

下面请听知情同意书的内容：

您是完全自愿参与本次试验的，本次试验没有经济上的补助。您提供的信息都是匿名的，会被严格保密，且仅供本研究使用。我们不会采集任何可能透露您身份的个人信息。这份知情同意书将会由研究负责人董琳娜女士保存在一个安全的地方。根据欧洲的法律，这份文件在试验结束 5 年后会被销毁。您可以无理由地随时退出试验，这将不会对您产生任何消极的影响。

如果您愿意参与试验，请您在听完下面的话之后，说一句“是的”。如果您说了“不”，那意味着您不同意进行本次试验。

- 有人为我阅读过本试验的说明，我已经知晓。

是

否

- 我有机会就试验进行提问。

是

否

- 我同意参加本次试验。

是

否

3.12. Informed consent form for the reception study (English)

This study belongs to a PhD research that aims at investigating the possibility of offering audio description using synthetic voices instead of a human voice. This is an example of a synthetic voice: “Hello, my name is Liang and I will be your host today”

Advances in technology mean synthetic voices are improving quickly and the reason we are exploring this option is that this could be a way to produce more audio described films more quickly. However, the aim is not to replace human voices in audio description but to find out whether Chinese AD users accept synthetic voices in film audio description.

The test, which will last approximately 40 minutes, consists of two parts. First you will listen to one clip that lasts around 5 minutes. Then you will have to answer some questions about it. After that, you will listen to a second clip of a similar duration and will have to answer some more questions. The two clips chosen come from Ann Hui’s historical film *Our time will come*. This movie revolves around the resistance movement during Japan’s occupation of Hong Kong.

The researcher in charge of the study is Mrs. Irene Tor. You can contact her at 13162099635 and ask her for more information about the study and the results.

Now please listen to the consent form:

Your participation in the tests is absolutely voluntary, and there is no economic compensation. The information you provide will be used in the study, but it will remain anonymous. All information that you give will be treated in the strictest confidence and it will only be used for the purposes of this doctoral research. No personal data will be collected that permits your identification. This audio consent will be kept in a safe place by the main researcher and according to the European law will be destroyed after five years once the investigation has finished. You can discontinue your involvement in the

study at any time without prior justification. This shall have no repercussions or negative consequences of any sort for you.

If you are willing to participate, please confirm the following statements by saying “yes” at the end. If you say “no” it means you do not give your consent.

- I have been read and have understood the information given for this research;
- I have had the opportunity to ask questions about this research;
- I consent to take part in the research sessions.

Annex 4: Questionnaires

4.1. Questionnaire for the audio describers (Chinese)

形象定义|调查问卷

第一部分：个人资料概况

调查问卷开头部分包括 8 个关于您的个人资料信息的问题，在问题之后有一个文本框，您可以根据自己的意愿在其中写下您的补充意见。

1. 您来自中国的哪个省份？
2. 您的母语是什么？
 - 普通话
 - 中国方言（请您注明）
3. 您的年龄在哪个区间？
 - 20-30
 - 31-40
 - 41-50
 - 51-60
 - 61-70
 - 71-80
4. 您的性别是？
 - 男
 - 女
5. 您的最高学历是什么？如果您还在上学，请选择您已获得的最高学历。
 - 未接受教育
 - 小学
 - 中学
 - 专科
 - 本科

- 硕士
 - 博士
 - 其他, 请注明:
 - 不对此作答
6. 您的视力状况如何?
- 视力正常
 - 局部失明
 - 全盲的
7. 您有视觉记忆吗?
- 有
 - 没有
 - 其他, 请说明:
8. 您现在的身份是?
- 无障碍音频描述工作的解说员/编剧
 - 无障碍音频描述作品的听众
 - 提供无障碍音频描述作品的单位
 - 为视力障碍者提供服务的某组织成员
9. 您还有什么需要补充的吗? (选填)

第二部分: 编剧/解说员的工作

1. 您做编剧/解说员多久了 (在无障碍音频描述工作的任何领域, 如电影、电视、直播活动等)?
- 不到 1 年
 - 1-5 年
 - 6-10 年
 - 11-15 年
 - 16-20 年
 - 20 年以上

2. 在无障碍音频描述制作领域，您的工作定位是？（可多选）
- 正式工作（有报酬）
 - 半职业的工作（微薄的或象征性的报酬，比如以电影票的形式获得报酬）
 - 志愿工作（无报酬）
3. 您在接受过何种类型的培训？（可多选）
- 未接受过培训
 - 电影
 - 电视
 - 博物馆
 - 戏剧
 - 歌剧
 - 其他现场活动
 - 无障碍教材
 - 其他，请说明：
4. 您的无障碍音频描述工作通常是制作什么类型的作品？（可多选）
- 电影
 - 电视
 - 博物馆
 - 戏剧
 - 歌剧
 - 其他现场活动
 - 无障碍教材
 - 其他，请说明：
5. 您在工作中制作了多少时长的无障碍材料？
- 不到 50 个小时
 - 51-150 个小时
 - 151-300 个小时

- 300 个小时以上
 - 其他, 请说明:
6. 您在以下制作过程的哪些无障碍音频描述的工作阶段富有经验? (可多选)
- 写剧本或修改剧本
 - 翻译剧本
 - 机器翻译加后编辑
 - 解说剧本
 - 利用发音天赋为解说提供帮助
 - 使解说与原声接轨
 - 控制最终作品的质量 (如检验剧本、录音的质量等)
 - 其他, 请说明:
7. 您在编辑 / 解说无障碍音频描述作品时使用何种语言?
8. 您在创作无障碍音频描述作品时, 是个人工作还是团队合作?
- 总是一个人工作
 - 主要是个人工作
 - 有时候一个人工作, 有时候进行团队合作
 - 主要是团队合作
 - 总是团队合作
9. 您在创作无障碍音频描述作品时, 您跟视力障碍者合作的频率如何?
- 总是
 - 经常
 - 偶尔
 - 不常
 - 从不
10. 假如您从来没有与他们合作过, 您觉得与视力障碍者的合作会有益于创作吗?
- 会
 - 不会

我不知道

11. 您还有什么需要补充的吗？（选填）

第三部分:无障碍音频描述工作及职业圈的情况

1. 在您从事无障碍音频描述工作之前，您从事过其他职业吗？

有

没有

2. 如果您从事过其他职业，那么您从事的是什么职业？

翻译

影视翻译（比如，字幕、配音翻译）

记者

作家

老师

广播/电视评论人

主持人

演员

电视制作人

其他，请说明：

3. 目前您是无障碍音频描述的全职工作者，还是您同时从事其他工作？

我是全职工作者

我还有其他工作，具体是：

我还是学生

4. 您是如何进入无障碍音频描述这一领域的？（可多选）

通过一门无障碍音频描述的专业课的学习之后

通过一门影视翻译课程的学习之后

在一个无障碍音频描述公司实习之后

我在这个行业工作的朋友帮助我进入这一行

我在这个行业工作的家人帮助我进入这一行

- 我创立了自己的无障碍音频描述公司
 - 根据我的履历 / 作品集被选择
 - 其他, 请说明:
5. 为了解决无障碍音频描述过程中的具体问题, 您向其他解说员 / 编剧咨询的频率如何?
- 很频繁
 - 常常
 - 偶尔
 - 不常
 - 从不
6. 从其他无障碍作品制作中获得的解决办法, 对您的工作影响有多大?
- 非常大的影响
 - 比较大的影响
 - 有一点影响
 - 很小的影响
 - 未受到影响
7. 您是否知道有无障碍音频描述工作指南的存在?
- 是
 - 否
8. 如果您知道, 那是什么样的指南? (可多选)
- 国家的指南
 - 内部指南
 - 其他国家的指南
 - 我自己制作的指南
 - 我不清楚
9. 您觉得指南重要吗?
- 是
 - 否

- 在一定程度上重要（具体的是）
10. 每次您都会接受提供给您的无障碍音频描述的工作吗？即使超出了您的无障碍音频描述专业的范围，您也答应吗？
- 是的
 - 不是
 - 我没专业的范围
11. 您有足够的时间成功地完成对作品的音频描述吗？
- 总是
 - 常常
 - 偶尔
 - 很少
 - 从不
12. 对您来说，您的无障碍音频描述工作是一种艺术（可以改进的天生才能）还是一种工艺（可以学习到的工作，专业培训的结果）？
- 肯定是一种艺术
 - 是一种工艺，但更是一种艺术
 - 两者都有一点
 - 是一种艺术，但更是一种工艺
 - 肯定是一种工艺
 - 我不清楚
13. 您如何评价以的说法（1 - 强烈不赞同；2 - 不赞同；3 - 不知道；4 - 同意；5 - 非常同意）？
- 参与无障碍作品的制作过程是能获得声誉的
 - 一般公众都了解无障碍音频描述工作
 - 盲人用户了解无障碍音频描述
 - 制作无障碍音频描述的工作压力很大
 - 制作无障碍音频描述的工作要求极高
 - 制作无障碍音频描述能赚很多钱

- 制作无障碍音频描述的工作是令人满意的
- 制作无障碍音频描述的工作是很有创意性
- 制作无障碍音频描述的工作是有益于社会的

14. 您认为以下哪一份工作与您从事的无障碍音频描述工作最相近?

- 影视翻译
- 编剧
- 作者
- 艺术家
- 技术员
- 演员
- 主持人-评论员
- 其他, 请说明:

15. 您还有什么需要补充的吗? (选填)

第四部分: 教育背景与无障碍音频描述工作的培训

此部分包括 7 个问题。

1. 您的教育背景是? (可多选)

- 语言与语言学
- 文学
- 翻译
- 电影与电视
- 戏剧
- 表演学校
- 艺术 / 博物馆
- 心理学
- 大众传媒学
- 科学
- 计算机科学和信息技术

- 其他，请说明：
2. 您接受过无障碍音频描述方面的培训吗？
- 是
 - 没有
3. 如果您接受过培训，是从哪里接受的？（可多选）
- 工作室
 - 职业课程
 - 大学课程
 - 实习中
 - 内部培训（由公司/培训提供的）
 - 一对一课程讲解
 - 其他，请说明：
4. 完成培训之后您获得了证书吗？
- 是，具体是：
 - 否
5. 如果您获得了证书，无障碍音频描述工作的公司在提供给您工作时，有要求您出示该证书吗？
- 有
 - 没有您如何继续提升您的技能和能力？（可多选）
 - 积累这个领域的经验
 - 参与讲座、工作室活动等
 - 做研究（比如，收集作品信息，跟电影/戏剧导演或博物馆工作人员讨论等）
 - 内部培训（由公司/机构提供的）
 - 分析已有的无障碍音频描述作品（关注同行采取的解决方案）
 - 学习现有的材料（指南、论文、相关书籍等）
 - 现在不考虑这个
 - 其他，请说明：

6. 您教过其他人如何制作无障碍音频描述作品吗？

- 是的
- 没有

7. 您还有什么需要补充的吗？（选填）

第五部分：技能与才干

这是调查问卷的最后一部分，包括 5 个问题。我们想要了解您从专业角度出发，哪些技能、能力与活动是您认为这一领域最合适和最需要的。请根据其重要性对以下项目进行评价，并在空白处写下我们忽略的内容。

1. 请根据 1 到 5 的等级（1- 不重要，2- 有点重要，3- 既重要又不重要，4- 重要，5- 非常重要）对以下陈述进行评价。

为了提供高质量的音频描述，无障碍音频描述工作者应该能够：

- ...解决问题
- ...具有良好的交际和人际交往能力
- ...坚定自信，为保证音频描述设备的质量努力
- ...应对时间压力
- ...有效地组织工作
- ...知道何时寻求专家帮助
- ...积极探索，评估，并酌情纳入反馈
- ...与同事一起工作
- ...与盲人赞助者的团队合作
- ...快速工作直到截止日期
- ...即兴发挥

2. 请根据 1 到 5 的等级（1 - 不重要，2 - 有点重要，3 - 既重要又不重要，4 - 重要，5 - 非常重要）对以下陈述进行评价。

为了提供高质量的音频描述，无障碍音频描述工作者应该在以下领域具备扎实的理论知识与理解力：

- 视听文本和多模式（取决于您的专业领域，包括剧院符号学，电影研究，艺术和博物馆研究等）
- 无障碍媒体（标准，立法，指南，原则和适用场景，技术等）
- 音频描述历史，状态和适用场景（例如，博物馆音频描述，电影音频描述，现场活动音频描述等）
- 音频描述原则，指南和标准
- 目标群体：视力障碍型，用户感知和认知处理，残疾人需求
- 翻译研究和影视翻译
- 语言和语言学（例如，了解文本分析、文本衔接和连贯的原则；掌握文学手法，如使用比喻、隐喻和象征语言；处理不同水平的语言形式等
- 编剧
- 世界知识

3. 请根据 1 到 5 的等级（1- 不重要，2- 有点重要，3- 既重要又不重要，4- 重要，5- 非常重要）对以下陈述进行评价。

为了提供高质量的音频描述，无障碍音频描述工作者应该在以下领域具备扎实的技术知识和技能：

- 音频描述脚本编写和文本编辑
- 使用音频描述软件
- 音频描述设备技术
- 音频描述解说
- 音频描述录音
- 音频描述与原声混合

4. 请根据 1 到 5 的等级（1- 不重要，2- 有点重要，3- 既重要又不重要，4- 重要，5- 非常重要）对以下陈述进行评价。

为了提供高质量的音频描述，无障碍音频描述工作者应该能够：

- …明确地表达意思
- …选择重要的视觉信息
- …使用激发想像力的语言

- …编译音频介绍
- …为听众提供可以“看到”所描述的画面方式
- …为听众提供一种可以“理解”所描述的画面方式
- …精通母语
- …使用非歧义语言
- …使用适合该作品的语言
- …使用适合大众的语言

5. 请根据 1 到 5 的等级（1- 不重要，2- 有点重要，3- 既重要又不重要，4- 重要，5- 非常重要，6- 我不清楚）对以下陈述进行评价。

您在音频描述时遇到的最困难的方面是什么？

- 解决问题
- 与他人的沟通和合作（同事和盲人）
- 坚定自信，为保证音频描述设备的质量努力
- 应对时间压力
- 有效地安排工作
- 知道何时请求专家帮助
- 积极探索，评估和酌情纳入反馈
- 与同事一起工作
- 与盲人赞助者的团队合作
- 快速编写直到截止日期
- 即兴发挥（例如，直播音频描述）
- 简明地表达意思
- 选择重要的视觉信息
- 使用激发想象力的语言
- 选择最合适的措辞
- 编辑音频介绍
- 为听众提供可以“看到”所描述的画面方式
- 为听众提供可以“理解”所描述的画面方式

- 精通母语
 - 使用非歧义语言
 - 使用适合该作品的语言
 - 使用适合观众的语言
 - 音频描述脚本编写和文本编辑
 - 使用音频描述软件
 - 使用音频描述设备技术
 - 音频描述解说
 - 音频描述录音
 - 将音频描述与原声混合
6. 您还有什么需要补充的吗？（选填）

这是问卷的最后一部分。非常感谢您能够完成问卷并帮助我们进行研究。您可以登录该项目网站 www.adlabproject.eu/ 获取调查结果及更新内容。谢谢!

4.2. Questionnaire for the audio describers (English)

IO2: PROFILE DEFINITION | QUESTIONNAIRE

SECTION 1: DEMOGRAPHIC PROFILE

This opening section of the questionnaire includes 8 questions on your demographic profile followed by a text box in which you can write your comments if you wish to do so.

1. What province in China are you from?
2. What is your mother tongue?
 - Mandarin
 - Other Chinese dialects (please specify which one)
3. What is your age?
 - 20-30
 - 31-40

- 41-50
 - 51-60
 - 61-70
 - 71-80
4. What is your gender?
- Male
 - Female
5. What is the highest degree or level of school you have completed? If currently enrolled, highest degree received.
- No schooling completed
 - Primary
 - Secondary
 - Vocational
 - BA/BSc
 - MA/MSc
 - PhD
 - Other, please specify:
 - Prefer not to answer
6. How would you describe your sight condition?
- Fully sighted
 - Partially sighted
 - Totally blind
7. Do you have a visual memory?
- Yes
 - No
 - Other/comments
8. Are you currently?
- Audio describer
 - AD user
 - AD provider
 - Member of organisation that provides support for people with sight loss
9. Is there anything you would like to add? (Optional)

SECTION 2: YOUR ACTIVITY AS AN AUDIO DESCRIBER

1. How long have you been working as an audio describer (in any area of AD, e.g. film, TV, live events)?
 - Less than one year
 - 1-5 years
 - 6-10 years
 - 11-15 years
 - 16-20 years
 - over 20 years
2. Is your activity as an audio describer (multiple answers possible):
 - Professional work (paid)
 - Semiprofessional work (a small, token payment or payment in kind e.g. tickets)
 - Voluntary work (no payment)
3. What type of audio description were you trained in? (multiple answers possible)
 - I had no training
 - Film
 - TV
 - museum
 - theatre
 - opera
 - other live events
 - AD of teaching materials
 - other, please specify:
4. What type of products do you audio describe most often? (multiple answers possible)
 - film
 - TV
 - museum
 - theatre

- opera
 - other live events
 - AD of teaching materials
 - other, please specify:
5. How much AD material have you produced in your career?
- Less than 50 hours
 - 51-150 hours
 - 151-300 hours
 - Over 300 hours
 - Other/comment:
6. Which of the following stages of the production process do you have experience with? (multiple answers possible)
- Writing and/or revising the AD script
 - Translating AD scripts
 - Using machine translation with post-editing
 - Voicing the AD
 - Assisting at recording the AD with voice talents
 - Mixing the AD with the original soundtrack
 - Quality control of the final product (e.g. checking the script or recording or both)
 - Other, please specify:
7. In what language(s) do you prepare your audio descriptions?
8. When you prepare audio descriptions, do you work alone or in a team?
- Always alone
 - Mainly alone
 - Sometimes alone and sometimes in a team
 - Mainly in a team
 - Always in a team
9. How often do you cooperate with persons with visual impairments when preparing audio descriptions?
- always
 - often

- sometimes
 - rarely
 - never
10. If you have never cooperated with persons with visual impairments when preparing audio descriptions, do you think that cooperation would have a positive influence on your work?
11. Is there anything you would like to add? (Optional)

SECTION 3: THE STATUS OF AD AND THE PROFESSIONAL CIRCLE

1. Did you work in another profession before you became an audio describer?
- Yes
 - No
2. If yes, what profession?
- Translator
 - Audiovisual translator (e.g. subtitler, dubbing translator, etc.)
 - Journalist
 - Writer
 - Teacher
 - Radio/TV commentator
 - Presenter
 - Actor
 - TV producer
 - Other, please specify:
3. Do you currently work as an audio describer full time or do you have other jobs?
- I work as an audio describer full time
 - I have other jobs (specify)
 - I am still a student
4. How did you start working as an audio describer? (multiple answers possible)
- After a specialized course on AD
 - After a course on audiovisual translation
 - After an internship in an AD company

- Someone working in the field (acquaintance) helped me enter this profession
 - Someone working in the field (family) helped me enter this profession
 - I created my own company
 - I was selected based on my CV or showreel/portfolio
 - Other, please specify:
5. How often do you ask the opinion of other audio describers to overcome specific AD problems?
- very often
 - often
 - sometimes
 - rarely
 - never
6. How much do the solutions found in other audio descriptions influence your work?
- A lot
 - Quite a lot
 - Somewhat
 - Very little
 - Not at all
7. Are you aware of the existence of audio description guidelines?
- Yes
 - No
8. If so, what type of guidelines are they? (multiple answers possible)
- National guidelines
 - In-house guidelines
 - Guidelines from other countries
 - Guidelines I created
 - I don't know
9. Do you think guidelines are important?
- Yes
 - No
 - To an extent (please comment)

10. Have you always accepted all audio description jobs that have been offered to you, even if they fall outside your AD specialization?
- Yes
 - No
 - I do not have a specialization
11. Do you have enough time to satisfactorily audio describe the products that are commissioned to you?
- always
 - often
 - sometimes
 - rarely
 - never
12. In your opinion, is the work of the audio describer an “art” (an innate talent that can be refined on the job) or a “craft” (an activity you learn, the result of specialized training)?
- definitely an “art”
 - more of an “art” than a “craft”
 - it is a bit of both
 - more of a “craft” than an “art”
 - definitely a “craft”
 - I don’t know
13. How would you rate the following statements on a 1 to 5 scale (1 – strongly disagree; 2– disagree; 3 – undecided; 4 – agree; 5 – strongly agree)?
- The work of the audio describer is prestigious
 - The work of the audio describer is well known by the general public
 - The work of the audio describer is well known by blind users
 - The work of the audio describer is stressful
 - The work of the audio describer is demanding
 - The work of the audio describer is well paid
 - The work of the audio describer is satisfying
 - The work of the audio describer is creative
 - The work of the audio describer is socially useful

14. In your opinion, which professional figure does a describer most closely resemble?

- Audiovisual translator
- Scriptwriter
- Author
- Artist
- Technician
- Actor
- Presenter - commentator
- Other, please specify:

15. Is there anything you would like to add? (Optional)

SECTION 4: EDUCATIONAL BACKGROUND AND AD TRAINING

This section includes 7 questions.

1. What is your educational background? (multiple answers possible)

- Language and linguistics
- Literature
- Translation
- Film and TV Studies
- Theatre studies
- Acting school
- Arts and/or museum studies
- Psychology
- Journalism or media studies
- Science
- Computer science and IT
- Other (specify)

2. Have you received specific AD training?

- Yes
- No

3. If so, in what form? (multiple answers possible)
 - workshop
 - vocational course
 - university course
 - internship
 - in-house training (conducted at a company/institution, etc.)
 - one-to-one instruction
 - other, please specify:
4. Did you get a certificate after completing the training?
 - Yes > go to following question
 - No
5. If yes, have you ever been asked to show that certificate in order to get work?
 - Yes
 - No
6. How do you continue to improve your skills and competences? (multiple answers possible)
 - Experience in the field
 - Participation in conferences, workshops, etc.
 - Doing research (e.g. gathering information on product, discussion with film/theatre directors/producers or with museum staff and curators, etc.)
 - In-house training (conducted at a company/institution, etc.)
 - Analysis of existing ADs (focusing on the solutions adopted by colleagues)
 - Study of existing material (guidelines, academic articles, books on AD, etc.)
 - Not at the moment
 - Other, please specify:
7. Have you ever taught AD yourself?
 - Yes
 - No
8. Is there anything you would like to add? (Optional)

SECTION 5: SKILLS AND COMPETENCES

This is the last section of the questionnaire, and it includes 5 blocks. We are interested in your opinion and views on the type of skills, competences and activities you consider most appropriate and needed by professionals looking to work in this area. Please, rate the following items in terms of their significance, and tell us in the spaces provided about anything we missed.

1. On a scale from 1 to 5 (1 – of no importance, 2 – of minor importance, 3 – neither important nor unimportant, 4 – important, 5 – extremely important), rate the following statements. **To deliver a good quality audio description, an audio describer should be able to ...** (Soft skills)
 - ... solve problems
 - ... have good communicative and interpersonal skills
 - ... be assertive and fight for the quality of AD provision
 - ... cope with time pressure
 - ... organize work efficiently
 - ... know when to call for expert help
 - ... actively seek, evaluate, and if appropriate incorporate feedback
 - ... work in a team with colleagues
 - ... work in a team with blind patrons
 - ... work quickly to a deadline
 - ... improvise
2. On a scale from 1 to 5 (1 – of no importance, 2 – of minor importance, 3 – neither important nor unimportant, 4 – important, 5 – extremely important), rate the following statements. **To deliver a good quality audio description, an audio describer should possess solid theoretical knowledge and understanding in the following areas:**
 - Audiovisual texts and multimodality (depending on your area of expertise, this can include for instance theatre semiotics, film studies, arts and museum studies, etc.)
 - Media accessibility (standards, legislation, guidelines, principles and applicable scenario's, technologies, etc.)

- AD history, status, and applicable scenarios (e.g. museum AD, film AD, AD for live events, etc.)
 - AD principles, guidelines and standards
 - Target group: types of visual impairment, user perception and cognitive processing, disabled needs
 - Translation studies and audiovisual translation
 - Language and linguistics (e.g. knowing the principles of text analysis, text cohesion and coherence; handling literary devices such as the use of similes, metaphors and figurative language; coping with different levels of formality in language; etc.)
 - Scriptwriting
 - World knowledge
3. On a scale from 1 to 5 (1 – of no importance, 2 – of minor importance, 3 – neither important nor unimportant, 4 – important, 5 – extremely important), rate the following statements. **To deliver a good quality audio description, an audio describer should be able to possess solid technical knowledge and skills in the following areas:**
- AD script writing and textual editing
 - Use of AD software
 - Technology of AD provision
 - AD voicing
 - AD recording
 - Mixing AD with original sound
4. On a scale from 1 to 5 (1 – of no importance, 2 – of minor importance, 3 – neither important nor unimportant, 4 – important, 5 – extremely important), rate the following statements. **To deliver a good quality audio description, an audio describer should be able to** (textual and linguistic skills):
- ... express meaning succinctly
 - ... select significant visual information
 - ... use language that sparks the imagination
 - ... compile an audio introduction
 - ... provide the listener with a way of "seeing" what is described

- ... provide the listener with a way of "understanding" what is described
 - ... possess an excellent command of mother tongue
 - ... use non-ambiguous language
 - ... use language that is suited to the product
 - ... use language that is suited to the audience
5. **What are the most difficult aspects you find when you audio describe?** Rate the following items on a scale from 1 to 5 (1 – extremely difficult, 2 – difficult, 3 – neither easy nor difficult, 4 – easy; 5 very easy; 6– I don't know)
- solving problems
 - communicating and cooperating with others (peers and blind persons)
 - being assertive and fighting for the quality of AD provision
 - coping with time pressure
 - organizing work efficiently
 - knowing when to call for expert help
 - actively seeking, evaluating, and if appropriate incorporating feedback
 - working in a team with colleagues
 - working in a team with blind patrons
 - writing quickly to a deadline
 - improvising (e.g. for live ADs)
 - expressing meaning succinctly
 - selecting significant visual information
 - using language that sparks imagination
 - choosing the most appropriate wording
 - compiling an audio introduction
 - providing the listener with a way of "seeing" what is described
 - providing the listener with a way of "understanding" what is described
 - possessing excellent command of mother tongue
 - using non-ambiguous language
 - using language that is suited to the product
 - using language that is suited to the audience
 - AD script writing and textual editing

- using AD software
 - using technology of AD provision
 - AD voicing
 - AD recording
 - mixing AD with original sound
6. Is there anything you would like to add? (Optional)

This was the last section of the questionnaire. Thank you very much for completing it and helping us with our research. Results and updates will be available on the project website www.adlabproject.eu/

4.3. Questionnaire for the end users (Chinese)

A. 口述影像消费习惯

- a) 您一般观看以下哪种形式的口述影像? (可多选)
- 电影院中的口述影像 (无障碍电影)
 - 博物馆中的口述影像
 - 电视节目中的口述影像
 - 体育赛事中的口述影像
 - 中国戏剧中的口述影像
 - 教学材料中的口述影像
 - 其他, 请说明:
 - 不知道, 不予作答
- b) 基于您所在地区的具体情况, 您希望未来以下哪个领域能引入口述影像服务? (可多选)
- 电影院
 - 博物馆
 - 电视节目
 - 体育赛事
 - 戏剧

- 教学材料
 - 其他, 请说明:
 - 不知道, 不予作答
- c) 您通常在哪里观看电影口述影像?
- 电影院
 - 家里
 - 盲人协会中的无障碍设施
 - 其他, 请说明:
 - 不知道, 不予作答
- d) 请指明您观看电影口述影像的频率:
- 很频繁 (只要条件允许)
 - 常常
 - 偶尔
 - 甚少
 - 从不
 - 不知道, 不予作答
- B. 电影口述影像服务满意度
- a) 您对所接触到的电影口述影像的数量感到满意吗?
- 非常满意
 - 比较满意
 - 一般
 - 不满意
 - 非常不满意
 - 不知道, 不予作答
- b) 您对中国电影口述影像的质量感到满意吗?
- 非常满意
 - 比较满意
 - 一般

- 不满意
- 非常不满意
- 不知道, 不予作答

C. 电影口述影像使用偏好

a) 在电影口述影像中您更喜欢以下哪种声音?

- 男声
- 女声
- 无所谓
- 不知道, 不予作答

D. 合成语音使用经历

a) 您是否使用带有合成语音的电子设备, 比如手机和电脑?

- 是
- 否
- 不知道, 不予作答

b) 如果您在上题中回答了“是”, 请说明使用频率:

- 很频繁
- 常常
- 偶尔
- 甚少
- 从不
- 不知道, 不予作答

c) 请指出在电影口述影像中, 合成语音应用研究的必要性。

- 非常有必要
- 有必要
- 没有必要
- 完全没有必要
- 不知道, 不予作答

d) 在必要的情况下, 您希望优先研究以下哪种类型的电影? (可多选)

- 动作片
- 爱情片
- 动画片
- 喜剧片
- 都市片
- 励志片
- 恐怖片
- 历史片
- 文艺片
- 战争片
- 科幻片
- 其他, 请说明:
- 不知道, 不予作答

E. 个人资料概况

- a) 年龄:
- b) 请指出您已获得最高学历:
 - 未接受教育
 - 小学
 - 中学
 - 专科
 - 大学
- c) 母语:
 - 普通话
 - 中国方言(请您注明)
- d) 如果您选择了“方言”, 请说明您的中文普通话水平:
 - 低级
 - 中级
 - 高级

- e) 您的视力状况如何?
- 全盲
 - 半盲
 - 其他情况, 请指明:
 - 不知道, 不予作答
- f) 如果您属于全盲或半盲, 请说明该视力状况是先天还是后天所致?
- 先天
 - 后天
 - 不知道, 不予作答
- g) 职业:
- 无工作
 - 已退休
 - 学生
 - 上班族
 - 其他, 请说明:
 - 不知道, 不予作答

十分感谢您的参与!

如果您想获取更多调查信息, 请联系调查员董琳娜, 电话联系方式 13162099593, 邮箱联系方式 irene.tor@uab.cat/ sara.rovira@uab.cat/ helena.casas@uab.cat 。谢谢!

4.4. Questionnaire for the end users (English)

A. Audio description habits

- a) What fields do you consume AD in? (Multiple answers accepted)
- a. Films
 - b. Museums
 - c. TV

- d. Sports events
 - e. Chinese opera
 - f. Teaching materials
 - g. Others, please specify
 - h. Prefer not to say
- b) Where would you like AD to be available? (Multiple answers accepted)
- a. Films
 - b. Museums
 - c. TV
 - d. Sports events
 - e. Chinese opera
 - f. Teaching materials
 - g. Others, please specify
- c) Where do you usually watch audio described films?
- a. At the cinema
 - b. At home
 - c. In the facilities of an association
 - d. Other, please specify
 - e. I don't know, prefer not to say
- d) How often do you watch audio described films?
- a. Very often (every time I have the chance)
 - b. Often
 - c. Sometimes
 - d. Rarely
 - e. Never
 - f. I don't know, prefer not to say

B. Satisfaction towards audio described movies

- a) Overall, are you satisfied with the provision (quantity) of ADs in your country?
- a. Very satisfied
 - b. Satisfied
 - c. Neither satisfied nor dissatisfied
 - d. Dissatisfied
 - e. Very dissatisfied

- f. I don't know, prefer not to say
- b) Overall, are you satisfied with the quality of ADs in your country?
 - a. Very satisfied
 - b. Satisfied
 - c. Neither satisfied nor dissatisfied
 - d. Dissatisfied
 - e. Very dissatisfied
 - f. I don't know, prefer not to say

C. Preferences regarding audio described films

- a) What type of voice do you prefer films to be audio described with?
 - a. Male voice
 - b. Female voice
 - c. I don't mind
 - d. I don't know, prefer not to say

D. Experience with artificial voices

- a) Do you use any equipment that incorporates artificial voices, such as mobile phones and computers?
 - a. Yes
 - b. No
 - c. Prefer not to say
- b) In case you selected "yes", please specify the frequency of use:
 - a. Very often
 - b. Often
 - c. Sometimes
 - d. Rarely
 - e. Never
 - f. Prefer not to say
- c) Do you think it is necessary to investigate whether artificial voices can be used in audio described movies?
 - a. Very necessary
 - b. Necessary
 - c. Not necessary

- d. Not necessary at all
- e. Prefer not to say
- d) In case you think it is necessary, in what type of movies should we start our research with?
 - a. Action films
 - b. Romantic films
 - c. Animation films
 - d. Urban films
 - e. Inspirational films
 - f. Horror films
 - g. Historical films
 - h. Artistic films
 - i. War films
 - j. Science fiction films
 - k. Other, please specify
 - l. Prefer not to tell

E. Demographic data

- a) What is your age?
- b) What is the highest degree or level of school you have completed? If currently enrolled,
 - a. No schooling completed
 - b. Primary school
 - c. Secondary school
 - d. Vocational training
 - e. University
- c) Mother tongue:
 - a. Standard Chinese
 - b. Chinese dialect, please specify.
- d) If you selected “Chinese dialect”, please specify your level of Standard Chinese:
 - a. Low
 - b. Average
 - c. High
- e) How would you describe your sight condition?

- a. Blind
 - b. Partially-sighted
 - c. Others, please specify.
 - d. Prefer not to say.
- f) In case you are blind or partially-sighted, is your visual loss congenital or acquired?
- a. Congenital
 - b. Acquired
 - c. Prefer not to say
- g) Occupation:
- a. Unemployed
 - b. Retired
 - c. Student
 - d. Employed
 - e. Others, please specify
 - f. Prefer not to say

Thanks for your participation!

If you want to know more about our project, you can contact Irene Tor. Her phone number is 13162099593 and her email is irene.tor@uab.cat

4.5. Questionnaire distributed by Sound of Light (Chinese)

为了进一步做好社区无障碍电影工作，更好地服务广大视障人士，我们设计了这份问卷。您的反馈对无障碍电影工作的完善很重要，期待听到您最真切的心声，谢谢配合！

2018年5月

一、答卷人基本信息

1. 视障程度

A 一级 B 二级 C 三级 D 四级

2. 性别:

A 男 B 女

3. 年龄:

A 18-25 B 26-35 C 36-45 D 46-59 E 60 及以上

4. 学历:

A 初中及以下 B 高中或中专中职 C 大专 D 本科及以上

5. 工作状态:

A 学生 B 上班族和挂靠 C 退休 D 其他

6. 所在街镇:

区 街道 (镇)

二、问卷部分

1. 您是否知道供视障人士观赏的无障碍电影?

A 知道 B 不知道 (答“不知道”的直接回答第 3 题)

2. 您是通过什么渠道了解到无障碍电影的?

A 残联盲协 B 助残员 C 朋友熟人 D 媒体和网络

3. 您是否有观赏无障碍电影的需要?

A 有需要 B 无需要 (答“无需要”的不再回答下面的问题)

4. 无障碍电影对您的意义是 (可多选):

A 娱乐 B 社交 C 学习 D 其他 (请说明)

5. 您曾经观赏过无障碍电影吗?

A 观赏过 B 没观赏过 (回答“没观赏过”的直接回答第 10 题)

6. 您对在社区放映无障碍电影感觉满意吗?

A 很满意 B 满意 C 一般 D 不满意

7. 您去社区观影的出行方式是

A 乘公交地铁 B 步行 C 其他 (请补充说明)

8. 观赏无障碍电影的过程中, 有听不懂的地方吗?

A 几乎没有 B 偶尔会有 C 经常会有

9. 您希望通过以下哪种方式获取影片放映信息?

A 相关微信公众号 B 微信群 QQ 群 C 上门通知 D 其他 (请说明)

10. 您喜欢观赏什么类型的电影 (最多可选三类):

A 动作片 B 爱情片 C 动画片 D 喜剧片 E 励志片 F 都市片

G 恐怖片 H 历史片 I 文艺片 J 战争片 K 科幻片 L 其他 (请说明)

11. 您偏爱以下哪种电影:

A 经典老片 B 最近新片 C 都可以

12. 您对观看无障碍电影还有什么建议?

真诚感谢您对本次调研的大力配合!

答卷日期 2018 年 月 日

4.6. Questionnaire distributed by Sound of Light (English)

In order to improve the provision of our barrier-free movies, and in order to better serve persons with sight loss, we designed this questionnaire. Your feedback is crucial to improve barrier-free movies. We are looking forward to your comments. Thanks for your cooperation!

May 2018

A. Demographic data

1. Grade of visual impairment
 - a. First
 - b. Second
 - c. Third
 - d. Fourth
2. What is your sex?
 - a. Male
 - b. Female
3. What is your age?
 - a. 18-25
 - b. 26-35
 - c. 36-45
 - d. 46-59
 - e. 60 or above
4. What is your educational background?
 - a. No schooling or primary education
 - b. Secondary education or medium vocational training
 - c. Advanced vocational training
 - d. University education
5. Job
 - a. Student
 - b. *Guakao*⁴⁴
 - c. Retired
 - d. Others
6. Where do you live? District and street

B. Questionnaire

1. Are you aware of the existence of barrier-free movies?
 - a. Yes
 - b. No (In case you select this answer, you can go straight to question number 3)

⁴⁴ Phenomenon wherein companies employ disabled staff but they do not have to turn up for work. Such companies save money by not spending on accessibility facilities and the disabled earn a salary for doing nothing.

2. How did you get to know about free-barrier movies?
 - a. Through an association
 - b. Through assistive staff
 - c. Through relatives and friends
 - d. Through the media and the Internet
3. Do you have the need to watch barrier-free movies?
 - a. Yes
 - b. No (In case you select this answer, your questionnaire ends here)
4. Why do you watch barrier-free movies?
 - a. For entertainment purposes
 - b. For socialising purposes
 - c. For educational purposes
 - d. Others (please specify)
5. Have you ever watched barrier-free movies?
 - a. Yes
 - b. No In case you select this answer, you can go straight to question number 10)
6. Are you satisfied with the barrier-free movies that you watch?
 - a. Very satisfied
 - b. Satisfied
 - c. Neither satisfied nor unsatisfied
 - d. Unsatisfied
7. How do you get to the screening facilities?
 - a. By public transport
 - b. On foot
 - c. Others (please specify)
8. Are there any moments in which you don't understand barrier-free movies?
 - a. Often
 - b. Sometimes
 - c. Almost none
9. How would you like to receive barrier-free movies-related information?
 - a. Through Wechat official accounts
 - b. Through Wechat and QQ groups

- c. Through door-to-door notification
 - d. Others (please specify)
10. What type of movies do you like to watch? (you can choose a maximum of three)
- a. Action movies
 - b. Romantic movies
 - c. Animation movies
 - d. Comedies
 - e. Inspirational movies
 - f. Urban movies
 - g. Horror movies
 - h. Historical movies
 - i. Artistic movies
 - j. War movies
 - k. Science fiction movies
 - l. Others (please specify)
11. What movies do you prefer?
- a. Classic
 - b. New
 - c. Both
12. Do you have any suggestions?

Our sincere thanks to your cooperation!

Questionnaire completed on (date)

4.7. Questionnaire for the reception study (A) (Chinese)

现在请看第一个片段。

请根据刚才的片段回答下列问题。

- 1. 方母给茅盾和他的妻子带了几块糕饼?
 - a. 2 块
 - b. 3 块
 - c. 一块也没带

- d. 我不知道/我不想说
2. 厨房的笼子里装了什么？
- a. 一只兔子
 - b. 一只蜗牛
 - c. 什么也没装
 - d. 我不知道/我不想说
3. 方兰的裙子是什么颜色的？
- a. 蓝色
 - b. 绿色
 - c. 白色
 - d. 我不知道/我不想说

请根据您刚才听到的解说声音回答问题，在每句话之后给出一个分数)

- 1. 刚才您听到的声音自然吗？（0分代表非常不自然，10分代表很像人的声音）
- 2. 刚才您听到的声音好听吗？（0分代表非常难听，10分代表非常好听）
- 3. 刚才您听到的声音的语调怎么样，是否有抑扬顿挫？（0分代表非常差，10分代表非常好）
- 4. 刚才您听到的声音发音标准吗？（0分代表非常不标准，10分代表非常标准）
- 5. 刚才您听到的声音在应该停顿的地方是否停顿了？（0分代表从不在应该停顿的地方停顿过，10分代表总是在应该停顿的地方停顿）
- 6. 刚才您听到的声音对您来说好理解吗？（0分代表非常不好理解，10分代表不费什么力就能理解）

7. 您觉得有些话您听不懂吗? (0分代表都听不懂, 10分代表都能听懂)
8. 您觉得刚才的声音可以接受吗? (0分表示完全不能接受, 10分表示完全能接受)
9. 您对刚才听到的声音的总体评价是怎么样的? (0分表示非常差, 10分表示好极了)
10. 您愿意继续观看带有这种声音的电影吗? (0分表示完全不愿意, 10分表示当然愿意)

现在请看第二个片段。

请根据刚才的片段回答下列问题。

1. 方兰在她母亲门口留了什么?
 - a. 一袋大米
 - b. 几封信
 - c. 一些药
 - d. 我不知道/我不想说
2. 当阿四到的时候, 方母正在做什么?
 - a. 做饭
 - b. 找女儿
 - c. 刮鱼鳞
 - d. 我不知道/我不想说
3. 为什么方母替阿四去了药店?
 - a. 因为阿四正在睡觉
 - b. 因为阿四病了

- c. 因为阿四不想去
- d. 我不知道/我不想说

请根据您刚才听到的解说声音回答问题，在每句话之后给出一个分数。

1. 刚才您听到的声音自然吗？（0分代表非常不自然，10分代表很像人的声音）
2. 刚才您听到的声音好听吗？（0分代表非常难听，10分代表非常好听）
3. 刚才您听到的声音的语调怎么样，是否有抑扬顿挫？（0分代表非常差，10分代表非常好）
4. 刚才您听到的声音发音标准吗？（0分代表非常不标准，10分代表非常标准）
5. 刚才您听到的声音在应该停顿的地方是否停顿了？（0分代表从没有在应该停顿的地方停顿过，10分代表总是在应该停顿的地方停顿）
6. 刚才您听到的声音对您来说好理解吗？（0分代表非常不好理解，10分代表不费什么力就能理解）
7. 您觉得有些话您听不懂吗？（0分代表都听不懂，10分代表都能听懂）
8. 您觉得刚才的声音可以接受吗？（0分表示完全不能接受，10分表示完全能接受）
9. 您对刚才听到的声音的总体评价是怎么样的？（0分表示非常差，10分表示好极了）
10. 您愿意继续观看带有这种声音的电影吗？（0分表示完全不愿意，10分表示当然愿意）

4.8. Questionnaire for the reception study (A) (English)

Please watch the first clip.

1. Please answer the following questions about the clip you just watched.

- a. How many rice cakes does Fang Mu bring to Mao Dun and his wife? (for A clip viewers)
 - i. 2
 - ii. 3
 - iii. None
 - iv. I don't know/Prefer not to tell
 - b. What was inside the kitchen's cage? (for A clip viewers)
 - i. A rabbit
 - ii. A snail
 - iii. Nothing
 - iv. I don't know/Prefer not to tell
 - c. What colour was Fang Lan's skirt? (for A clip viewers)
 - i. Blue
 - ii. Green
 - iii. White
 - iv. I don't know/Prefer not to tell
2. Please answer the following questions about the voice you just heard. Please give a mark for each of the following statements according to the voice you just heard.
- a. How natural was the voice you just heard? (0 means "very unnatural" and 10 means "like a human voice")
 - b. How pleasant was the voice you just heard? (0 means "very unpleasant" and 10 means "very pleasant")
 - c. How would you rate the intonation of the voice? (0 means "very bad" and 10 means "very good")
 - d. How would you rate the voice's pronunciation? (0 means "very bad" and 10 means "very good")
 - e. Does the voice make the pauses when they are needed? (0 means "never" and 10 means "always")
 - f. How easy was it for you to understand the voice? (0 means "very difficult" and 10 means "I had to make no effort at all")
 - g. Did you find certain words hard to understand? (0 means "never" and 10 means "all the time")
 - h. Do you find this voice acceptable? (0 means "totally unacceptable" and 10 means "totally acceptable")

- i. What is your overall impression of the voice you just heard? (0 means “very bad” and 10 means “excellent”)
3. In a scale from 0 to 10, would you like to keep on watching the movie with this voice? (0 means “not at all” and 10 means “for sure”)

4.9. Questionnaire for the reception study (B) (Chinese)

现在请看第一个片段。

请根据刚才的片段回答下列问题。

1. 方兰在她母亲门口留了什么?
 - a. 一袋大米
 - b. 几封信
 - c. 一些药
 - d. 我不知道/我不想说
2. 当阿四到的时候, 方母正在做什么?
 - a. 做饭
 - b. 找女儿
 - c. 刮鱼鳞
 - d. 我不知道/我不想说
3. 为什么方母替阿四去了药店?
 - a. 因为阿四正在睡觉
 - b. 因为阿四病了
 - c. 因为阿四不想去
 - d. 我不知道/我不想说

请根据您刚才听到的解说声音回答问题，在每句话之后给出一个分数。

1. 刚才您听到的声音自然吗？（0分代表非常不自然，10分代表很像人的声音）
2. 刚才您听到的声音好听吗？（0分代表非常难听，10分代表非常好听）
3. 刚才您听到的声音的语调怎么样，是否有抑扬顿挫？（0分代表非常差，10分代表非常好）
4. 刚才您听到的声音发音标准吗？（0分代表非常不标准，10分代表非常标准）
5. 刚才您听到的声音在应该停顿的地方是否停顿了？（0分代表从没有在应该停顿的地方停顿过，10分代表总是在应该停顿的地方停顿）
6. 刚才您听到的声音对您来说好理解吗？（0分代表非常不好理解，10分代表不费什么力就能理解）
7. 您觉得有些话您听不懂吗？（0分代表都听不懂，10分代表都能听懂）
8. 您觉得刚才的声音可以接受吗？（0分表示完全不能接受，10分表示完全能接受）
9. 您对刚才听到的声音的总体评价是怎么样的？（0分表示非常差，10分表示好极了）
10. 您愿意继续观看带有这种声音的电影吗？（0分表示完全不愿意，10分表示当然愿意）

现在请看第二个片段。

请根据刚才的片段回答下列问题。

1. 方母给茅盾和他的妻子带了几块糕饼？
 - a. 2块
 - b. 3块
 - c. 一块也没带

- d. 我不知道/我不想说
2. 厨房的笼子里装了什么？
- a. 一只兔子
 - b. 一只蜗牛
 - c. 什么也没装
 - d. 我不知道/我不想说
3. 方兰的裙子是什么颜色的？
- a. 蓝色
 - b. 绿色
 - c. 白色
 - d. 我不知道/我不想说

请根据您刚才听到的解说声音回答问题，在每句话之后给出一个分数。

- 1. 刚才您听到的声音自然吗？（0分代表非常不自然，10分代表很像人的声音）
- 2. 刚才您听到的声音好听吗？（0分代表非常难听，10分代表非常好听）
- 3. 刚才您听到的声音的语调怎么样，是否有抑扬顿挫？（0分代表非常差，10分代表非常好）
- 4. 刚才您听到的声音发音标准吗？（0分代表非常不标准，10分代表非常标准）
- 5. 刚才您听到的声音在应该停顿的地方是否停顿了？（0分代表从没有在应该停顿的地方停顿过，10分代表总是在应该停顿的地方停顿）
- 6. 刚才您听到的声音对您来说好理解吗？（0分代表非常不好理解，10分代表不费什么力就能理解）
- 7. 您觉得有些话您听不懂吗？（0分代表都听不懂，10分代表都能听懂）

8. 您觉得刚才的声音可以接受吗? (0 分表示完全不能接受, 10 分表示完全能接受)
9. 您对刚才听到的声音的总体评价是怎么样的? (0 分表示非常差, 10 分表示好极了)
10. 您愿意继续观看带有这种声音的电影吗? (0 分表示完全不愿意, 10 分表示当然愿意)

4.10. Questionnaire for the reception study (B) (English)

Please watch the first clip.

1. Please answer the following questions about the clip you just watched.
 - a. What does Fang Mu leave at her mother's door? (for B clip viewers)
 - i. A bag of rice
 - ii. Some letters
 - iii. Some medicine
 - iv. I don't know/Prefer not to tell
 - b. When A Si arrives, what is Fang Mu doing? (for B clip viewers)
 - i. Cooking
 - ii. Looking for her daughter
 - iii. Scaling a fish
 - iv. I don't know/Prefer not to tell
 - c. Why does Fang Mu go to the pharmacy instead of A Si? (for B clip viewers)
 - i. Because A Si is sleeping
 - ii. Because A Si is sick
 - iii. Because A Si does not want to go
 - iv. I don't know/Prefer not to tell
2. Please answer the following questions about the voice you just heard. Please give a mark for each of the following statements according to the voice you just heard.
 - a. How natural was the voice you just heard? (0 means "very unnatural" and 10 means "like a human voice")
 - b. How pleasant was the voice you just heard? (0 means "very unpleasant" and 10 means "very pleasant")

- c. How would you rate the intonation of the voice? (0 means “very bad” and 10 means “very good”)
 - d. How would you rate the voice’s pronunciation? (0 means “very bad” and 10 means “very good”)
 - e. Does the voice make the pauses when they are needed? (0 means “never” and 10 means “always”)
 - f. How easy was it for you to understand the voice? (0 means “very difficult” and 10 means “I had to make no effort at all”)
 - g. Did you find certain words hard to understand? (0 means “never” and 10 means “all the time”)
 - h. Do you find this voice acceptable? (0 means “totally unacceptable” and 10 means “totally acceptable”)
 - i. What is your overall impression of the voice you just heard? (0 means “very bad” and 10 means “excellent”)
3. In a scale from 0 to 10, would you like to keep on watching the movie with this voice? (0 means “not at all” and 10 means “for sure”)

4.11. Post-questionnaire for the reception study (Chinese)

请回答下列问题。

1. 您更喜欢哪个声音?
 - a. 第一个
 - b. 第二个
 - c. 两个都喜欢
 - d. 我不知道/我不想说
2. 总的来说，您要是能选择的话，您愿意在无障碍电影中听到哪种声音?
 - a. 人声
 - b. 机器人语音

- c. 取决于电影
 - d. 我不在乎
 - e. 我不确定/我不想说
3. 如果您上一题回答了“取决于电影”，那么您觉得哪种电影可以使用机器人语音呢？您可以选择多项。
- a. 纪录片
 - b. 历史片
 - c. 爱情片
 - d. 武术片
 - e. 喜剧
 - f. 悲剧
 - g. 惊悚片
 - h. 动作片
 - i. 奇幻片
 - j. 动画片
 - k. 其他类型（请具体说明）
4. 如果无障碍电影没法使用人声的话，您能接受临时使用机器人语音吗？
- a. 能
 - b. 不能
 - c. 我不知道/我不想说
5. 无障碍电影中，您能接受用机器人语音来永久替代人声吗？
- a. 能

- b. 不能
 - c. 我不知道/我不想说
6. 对此您有什么建议吗?
7. 您的性别是?
- a. 男性
 - b. 女性
 - c. 我不想说
8. 您的出生年份是?
9. 您的最高学历是?
- a. 我没有上过学
 - b. 小学
 - c. 中学
 - d. 大学
 - e. 我不想说
10. 您曾经接触过机器人语音吗?
- a. 是的
 - b. 没有
 - c. 我不想说
11. 您经常使用机器人语音吗?
- a. 很频繁地使用
 - b. 经常使用
 - c. 有时使用

d. 基本不用

e. 从没用过

f. 我不想说

12. 如果您用过机器人语音，您在什么情况下使用它？

13. 您经常看无障碍电影吗？

a. 总看

b. 经常看

c. 有时看

d. 基本不看

e. 从不看

f. 我不想说

最后，您有什么想说的吗？

感谢您的配合！

4.12. Post-questionnaire for the reception study (English)

1. Which voice did you like the most?

a. The first one

b. The second one

c. I liked both

d. I don't know/Prefer not to tell

2. In general, if you had a choice, which AD voice would you prefer?

a. A human voice

b. An artificial voice

c. Depends on the film

d. I don't care

e. Not sure

3. In case you replied “depends on the film”, for which kind of films would you find synthetic voices acceptable? You can choose more than one option.
 - a. Documentaries
 - b. Historical films
 - c. Romantic films
 - d. Martial arts films
 - e. Comedies
 - f. Tragic films
 - g. Horror films
 - h. Action films
 - i. Fantasy films
 - j. Animation
 - k. Others (please, specify):
4. Would you accept TTS AD as an interim solution, in case a human voice is not available to voice the AD?
 - a. Yes
 - b. No
 - c. Don't know/Prefer not to tell
5. Would you accept TTS AD as a permanent solution, as an alternative to a human voice?
 - a. Yes
 - b. No
 - c. Don't know/Prefer not to tell
6. Do you have any comments?
7. Sex
 - a. Male
 - b. Female
 - c. Prefer not to tell
8. Year of birth
9. Highest level of education received
 - a. I did not go to school
 - b. Primary education
 - c. Secondary education
 - d. Higher education

- e. Prefer not to tell
10. Do you have experience with TTS?
- a. Yes
 - b. No
 - c. Prefer not to tell
11. How often do you use TTS?
- a. Very often
 - b. Quite often
 - c. Sometimes
 - d. Rarely
 - e. Never
 - f. Prefer not to tell
12. In case you use it, in what context do you use TTS?
13. How often do you watch audio described films?
- a. Very often
 - b. Quite often
 - c. Sometimes
 - d. Rarely
 - e. Never
 - f. Prefer not to tell
14. Are there any final remarks that you would like to make?
- Thanks for your collaboration!

Annex 5: AD Scripts

5.1. Clip A's script

05:14 1941 年冬，香港沦陷。

05:17 出现一幢老屋，一个中年妇女端着瓷盘进屋，拉开橱柜抽屉拿出一个巴掌大的纸包；她就是女主人公方兰的母亲。

05:26 只见她小心翼翼地拆开纸包，里面是几只白色的糕饼；方母拿了两只糕饼放进碗里，拿到第三只时停了停，又放回纸包重新包上；

05:41 椭圆形梳妆镜，镜子里方母理了理发髻，又侧头检视一眼，而后端起盘子走出屋子。

05:53 方母攀着扶手缓步上楼，边走边招呼：

06:08 【快】方母倒茶。

06:14 【快快】沈先生叫沈德鸿，就是作家茅盾，他正在看书。

06:50 沈先生一声咳嗽，沈太太忙转话头，

07:09 方母敷衍着笑笑，

07:36 方母怏怏地端起糕饼，

07:40 方母下楼去了。

07:42 【快】沈太太转头看丈夫。

07:49 茅盾先生说完，继续低头看书；沈太太拿起杯子，喝了口；她又把另外半杯茶水倒回茶壶，闭上眼轻轻吁了一口气。

08:02 厨房里，方母伸手在米缸舀米，米缸已经见底，只舀出小半罐头。她把罐头里的米抖回去一些，再抖回去一些，然后把米倒进桌上的砂锅，盖上米缸。

08:18 方母起身，盖上砂锅锅盖，拿起一把菜刀，试了试刀刃，移步走到半人高的笼子旁，弯腰一看，笼子里空了；

08:34 检查笼子提闸，没有问题呀，

08:37 方母躬着身子，在厨房四处搜寻。她无奈地站起身，一扔菜刀，仰头望天花板。镜头一转，

08:54 一个身穿绿格裙子的姑娘抱着兔子，向草坪边树林走去。

09:03 近景镜头，一只蜗牛在草丛里吐泡沫。一男青年拿起蜗牛摆弄；姑娘来到他身边蹲下。

09:36 兔子在树林里吃草。

09:41 姑娘叫方兰，男青年叫李锦荣，是她的男友。两人并肩坐在草地上，李锦荣掏出个纸包，冲方兰一笑打开，方兰见是两块蛋糕。

09:54 方兰一笑伸手去拿，李锦荣躲开，

10:03 方兰虔诚地双手合十，闭上眼许愿。李锦荣含情凝视。

5.2. Clip B's script

01: 21: 27 镜头回到方家。画面中是一面镜子，镜子里方母左手托腮在发呆，她盘坐在床上，微闭双眼在打盹儿，但又马上惊醒了；

01: 21: 43 方母来到阳台上，看着小雨淅淅沥沥。稍后，她又踮起脚尖朝楼下张望，

01: 21: 55 好像看到了什么，方母蹒跚着走到楼梯口，只见方兰的身影在家门口一闪。方兰警惕地看看左右，把一小袋米放在门口；

01: 22: 08 方母捧起米袋往外追出几步，

01: 22: 13 被母亲扳住肩头，方兰这才转身；

01: 22: 27 方母高兴地拽着女儿往家走；

01: 22: 32 【慢】进屋后，方母捧起米袋中的米，放在鼻下闻闻，又让米粒从只缝间漏回米袋，还细心捡起桌上的米粒；

01: 22: 40 方兰看着母亲，有点心酸。她拿起斗笠和提篮，

01: 23: 09 【慢】尽管方兰摇头，但是方母还是拎着篮子出门了。方兰站在阳台上，俯望着苍老的母亲慢慢地走远。

01: 23: 22 画面切换，宪兵司令部油印室，胖胖的日本军曹站在窗边抽烟；

01: 23: 27 镜头转到废纸篓，里面有张印坏了的讨伐游击队军事部署图；

01: 23: 34 李锦荣一边操作油印机，一边迅速弯腰捡起那张军事部署图，揉成一团塞进裤兜，扫了一眼胖军曹，若无其事地继续工作；

01: 23: 50 方母在家中刮鱼鳞，阿四上楼来：

01: 24: 17 方母拿过咸鱼快步去厨房，阿四坐下擦汗。不一会儿方母端着粥出来，

01: 24: 48 饿急了的阿四双手捧着碗大口大口喝粥。微风吹拂的窗帘轻轻飘动；

01: 24: 58 阿四喝完粥坐在那儿打盹儿，后来干脆向后一靠，沉沉地睡去了；

01: 25: 06 方母让阿四在长椅上躺下，盖上薄被；她在椅边坐下，注视着这个年纪轻轻就投身抗日的姑娘，

01: 25: 16 【慢】方母起身拎起阿四的竹篮，拨开表层的草药查看了一下，又低头瞟一眼熟睡的阿四，拎着竹篮轻手轻脚下了楼。熟睡的阿四浑然不觉；

01: 25: 37 方母提着篮子，有点紧张地赶往联络点伍记药店；

01: 25: 45 药店老板娘伍姐在前台称药材；方母进门，

01: 26: 00 方母回身关上药店的大门

Annex 6: Interviews

6.1. Interview 1 (Confucius Institute)

上外“孔子新汉学计划”博士生董琳娜：从西班牙到中国，从翻译三毛作品到为残障人士发声

From: SISUOXI 上外孔院工作处 Today

三毛翻译家、西班牙巴塞罗那自治大学汉语老师、“加泰罗尼亚跨媒体”研究小组成员，这些头衔都用于称呼一个人——上外“孔子新汉学计划”博士生董琳娜（Irene Tor Carroggio）。是她，将三毛的文字带进了西班牙，使西语国家的人们有机会了解这位华人女作家的西班牙故事。也是她，致力于在中国研究推广“无障碍电影”，希望能够让中国视障人群拥有更加丰富的精神文化生活。在文明交流互鉴的路上，她从未停止过脚步。



无障碍电影是指在普通电影原有的声画基础上，增加故事旁白、人物对话，采用专业设备制作完成的、供有视力障碍或听力障碍的人们欣赏的电影。

一个选择，与中国相遇

2008年，还是高中生的董琳娜已经完成了英语和法语两门外语的官方考试。她的母亲建议，既然喜欢语言，为什么不试试学习第三门外语。“我选择了汉语，因为那时欧洲很多人认为，**汉语会是一门非常有前景的语言，中国正在急速发展，将来用中文交流必不可缺。**”抱着试试的态度，董琳娜开始了汉语的学习。但她没有想到，她的这个选择，改变了往后的人生道路。



随着汉语水平的不断提高，董琳娜发现自己越来越喜欢这门语言，她放弃了原本想要学医的想法，而是准备在大学攻读与汉语有关的学位。随后她考入了西班牙巴塞罗那自治大学的翻译专业，学习英西、汉西翻译。汉语学习并非易事，大学四年，她同班的同学从20人锐减到4人。但是董琳娜没有放弃，她希望自己能¹在汉西翻译领域走得更远。



董琳娜在柏林“语言与传媒”国际学术研讨会上发言

一位作家，指引她寻梦中国

在大学期间，董琳娜从一位中国留学生朋友的口中得知了作家三毛，通过交流，她越来越被这位女作家的传奇一生所吸引。为了更加了解三毛，她开始查阅资料，但却非常遗憾地发现，尽管三毛与西班牙有如此多的故事，西班牙人却对她一无所知。作为一名汉西翻译专业的学生，董琳娜下定决心，要让三毛走进西班牙。然而在西班牙，汉语翻译的资源十分有限，为了进一步提高自己的翻译水平，获取更为一手的翻译材料，董琳娜决定来中国留学。



董琳娜参加第二届“三毛散文奖”新闻发布会

一位翻译家，永不停止的理想追寻

大学毕业后，董琳娜先后到哈尔滨和北京进修汉语。与此同时，也没有放弃她的翻译工作，她一边用西语翻译三毛作品，一边尝试联系西班牙出版社询问意向。但是由于西班牙出版社对三毛所知甚少，董琳娜在出版社屡屡碰壁。2015年，经过了几年的努力，当时在上海读硕士的董琳娜终于收到了回音。西班牙RATA出版社联系到她，表示愿意将这位“由心写作的作家”推向西班牙。2016年，她的第一本译著《撒哈拉的故事》成功出版，在当地获得了不小的反响。乘胜追击，董琳娜的第二本译著《梦中的橄榄树》在2017年面世。



董琳娜的三毛译著

一名博士生，与上外结缘

硕士毕业后，董琳娜并没有满足于目前的成就，获得了母校巴塞罗那自治大学奖学金的她，回到了西班牙攻读“翻译与跨文化研究”专业的博士学位。作为她博士项目的一部分，董琳娜加入了“加泰罗尼亚跨媒体”研究小组，进行“媒体可及性”（Media Accessibility）研究。她将自己的汉西翻译专长与研究相结合，调查中国“无障碍电影”发展现状。



董琳娜与“无障碍电影”

“就像让三毛走进西班牙一样，我也想让‘无障碍电影’走进中国”。

无障碍电影，这个乍听起来有些陌生的名词，其实是指在普通电影原有的声画基础上，增加故事旁白、人物对话，采用专业设备制作完成的、供有视力障碍或听力障碍的人们欣赏的电影。尽管“无障碍电影”在欧洲和美国已经开始了几年的时间，但在世界上残疾人数量最多的中国，研究仍属于萌芽期。网络资源毕竟有限，为了实地考察中国情况，在西班牙完成了前期准备的董琳娜，打算回到中国开展下一步调研。不过这时她却发了愁，到了中国，她要如何联系这些残疾人群体？如何有效地推进调研呢？

偶然间，她从孔子学院总部官网查询到了“孔子新汉学计划”，了解到其“中外合作培养博士项目”可以资助外国大学在读博士生来华联合研究。董琳娜想，这可以成为她研究的突破口。她开始在参与“孔子新汉学计划”的中国十几所著名高校里找寻适合自己的导师，上海外国语大学英语学院的肖维青老师的研究方向和简介映入眼帘，“肖老师所研究的视听翻译正是我要找的方向！”经过层层选拔，董琳娜带着对研究的期望来到了上外。



董琳娜与导师肖维青教授

“这是我做过的最无悔的选择，肖老师太棒了！”肖老师不仅在论文上给予董琳娜指导，同时也给了她很多机会，让她能够走进课堂，向中国学生介绍“无障碍电影”。这也是董琳娜所希望的。她很高兴自己可以让更多人了解“无障碍电影”。或许明天，会有更多人加入她的行列，和她一起为残疾人争取更多的权益。



董琳娜在课堂上为中国学生介绍“无障碍电影”

现在，董琳娜奔忙于电影院、残联、志愿者协会之间，调查上海各区制作及播映“无障碍电影”的基本情况。她也试图为“无障碍电影”找寻一个更广、更便利的发展方向。目前“无障碍电影”主要借助志愿者协会录播辅助音频，董琳娜则考虑使用人工合成声音，以便将其推广至中国欠发达地区，尤其是农村，让更多的残疾人拥有丰富的精神文化生活。



董琳娜、肖维青教授与志愿者协会的韩颖老师

为了实现这一目标，董琳娜深入到残疾人家庭当中，用一段人工合成音频实验视障人士的可接受度。她的努力也吸引了上海电视台外语频道ICS的注意，在接受专访时表示：“我很期待实验结果，希望能够通过我的研究成果推动中国“无障碍电影”的发展，那样我中西文化‘信使’的工作也就完成了。”



董琳琳就“无障碍电影”接受上海电视台外语频道ICS采访

采访视频链接：<https://www.shanghaieye.com.cn/an-expats-push-for-a-more-accessible-shanghai/?from=groupmessage&isappinstalled=0>

这项实验结束后，董琳琳将回到西班牙完成她的博士论文。而关于未来，董琳琳给了自己许多种可能性。不过可以确定的是，她的未来，一定和中国有关。最近董琳琳已经收到了小说《白蛇传》的翻译邀请，博士毕业后，她将会继续自己的翻译事业。此外，她也会拓展在“无障碍电影”领域的研究，深化自己的中西文明互鉴之路。



董琳琳在上外

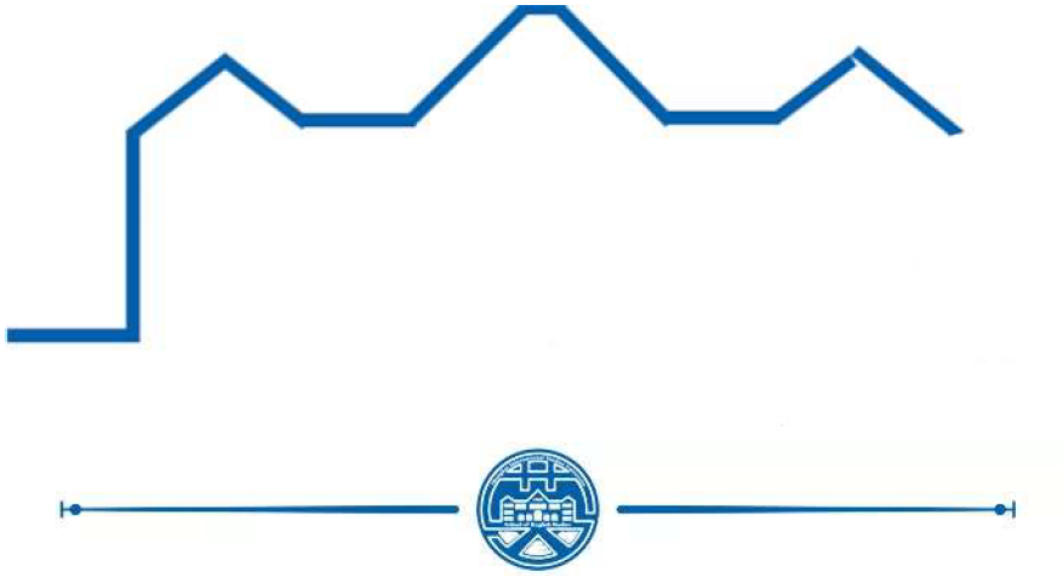
董琳琳的经历只是上海外国语大学“孔子新汉学计划”博士生的一个缩影。近年来，已有来自于埃及、越南、吉尔吉斯斯坦、韩国等国的博士生来到上外，进行中阿关系、外交学等领域的研究。上海外国语大学孔院工作处积极组织各学院之间的“孔子新汉学计划”博士生和导师座谈，从生活和学习各方面对他们进行帮助。今后，上外孔院工作处也将在孔子学院总部指导下，积极帮助各国青年学者研究中国，理解中国，书写“美美与共”新篇章。



6.2. Interview 2 (Shanghai International Studies University)

董琳娜：光影有声 进无止境

上外英院分团委学生会 Today



董琳娜：光影有声 进无止境



人物简介

*

董琳娜Irene Tor Carroggio，来自西班牙的联合培养博士，是上海外国语大学来自欧洲的第一位参与孔子新汉学计划的博士生。她从十七岁开始学习汉语，同时也能熟练运用英语，超强的多语种的学习能力在帮助她专业学习的同时，也让她体会到不同语言、不同文化的魅力。在导师的带领下，她投身于公益性的学术课题“无障碍影视翻译”，公益背后，是青年人想用自己的光点亮世界的赤诚之心。

求学异国 文化借行

董琳娜的家乡在西班牙巴塞罗那，她的汉语和英语也十分流利。说起董琳娜与中文和中国的缘分，始于她十七岁出于兴趣，在校内自主学习中文。“我很喜欢汉语，它很动听。最喜欢中国文化中对于友谊的看法。”提起汉语和中国文化，董琳娜这样说。“孔子新汉学计划”是由孔子学院总部设计的重要发展规划，旨在帮助世界各国青年了解中国和中华文化，繁荣汉学研究，增进中国与各国人民之间的友好关系，通过这一计划，来到中国留学之后，董琳娜也走过了许多地方，哈尔滨、北京和上海都留下过她求学的足迹，也在这个陌生的国度以语言为纽带，结交到了许多志同道合的友人，也是用语言为路标，指引出更广阔的未来蓝图。“起初为了有更好的语言环境来练习汉语，我选择了欧美留学生较少的哈尔滨，也是在那里，我的中文有了很大的进步。后来又到过北京，第三次我来中国求学，我选择到上海来攻读联合培养博士学位，并开始研究残障人士影视翻译。”从前的兴趣变成了现在学习、工作的语言之一，董琳娜对中文和中国始终怀着初见时的热情：“我中文名字的姓是我根据自己西班牙语名字的相似发音起的，汉语的每个字都有不同的音调，这对我来说有点难，不过每门语言都有它的魅力。”



虽然在中国时间不长，但课题研究之余，董琳娜的生活也并不单调。不久前，她受邀担任英语学院配音大赛的评委。比赛前，她以为参赛的都是专业的学生（比如表演专业，配音专业），但逐渐意识到这只是同学们的业余爱好，“我觉得特别新奇也很惊喜，我在读书的时候没有过这种活动，这次的活动比赛选手的水平都很高，”同时她也从自己研究课题的角度给予了同学们一些建议，“配音不仅要关注影视中对话的声音语言，说话人的肢体动作和其他行为语言也属于配音应该关注的内容。”在采访时，她了解到短视频配音是当下中国的潮流，欣喜地下载体验了一些配音软件，准备作为课题的辅助工具。



心系公益 为爱发声

董琳娜在硕士期间的专业是影视翻译，得到攻读博士学位的奖学金后，她决定将自己的学术课题与一直抱有浓厚兴趣的中国文化结合起来，最终她离开哈尔滨来到上海，并确定了“无障碍电影”这一课题。她谈到，“最初我选择来到上海，是因为大城市在‘无障碍电影’方面的普及度较高。来到这里之后我发现，这方面的实践和研究虽然仍较少，却有着光明的发展前景和广阔的传播平台。这也给了我很多鼓舞，我想在这方面多实践，为残障人士贡献自己的力量。”

无障碍电影是为方便残障人士观看而经过加工处理的电影，通常是通过重新剪辑增补大量配音解说的方式，让视力障碍者能够完整解读整部电影，体会观影的乐趣。想要达到好的解说和配音效果，就必须研究盲人的诉求，了解他们期待解说什么，同时也要考虑如何呈现无声的画面、合成音是否可以代替人声等。为此，董琳娜先后调研了无障碍电影在中国城市的普及情况，了解解说员的状况。“在中国，无障碍电影的解说员基本都有自己的工作，业余时间志愿无偿为电影解说，但在国外很多地方，无障碍电影解说已经发展成为正式的职业，中国在这一点上或许还有很长的路要走。”



目前在上海，无障碍电影的呈现形式是在每个月的最后一个周四，某一社区电影院会组织盲人在家人的陪同下一同免费观影，另外电影的录音带也会在指定渠道发售，给残障人士提供更多选择。董琳娜常作为志愿者进行现场口述，“这非常有挑战性，也意义非凡”。为促进无障碍电影解说行业的进步发展，董琳娜的导师肖维青教授邀请韩颖（全国自强模范，同时也是上海第一位获得英语中级口译证书的盲人）来到翻译课堂，听她和同学们对电影片段进行口述。韩颖老师比较了翻译和解说的区别，并提出了改进意见，这让董琳娜对无障碍电影解说有了更清晰的理解和更精准的把握，体会到无障碍电影解说行业发展的必要性和广阔前景。董琳娜从实践中出发，为“无障碍电影”理论不断添砖加瓦，逐渐摸索着适合中国国情的无障碍电影发展方式。



教学相长 步履不停

在上海外国语大学攻读学位的这段时间里，董琳娜既是心系公益认真研习学术课题的博士生，也是在给本科生授课过程中不断进行自我提升的助教。

上外是上海唯数不多招收影视翻译方向博士研究生的大学，因此董琳娜硕士毕业后选择上外，而这一选择，也造就了她和肖维青教授的师生缘。提起自己的导师，董琳娜重复很多遍的词语是“感谢”。“关于导师，我要说的就是‘感谢’，肖老师在课题研究上给予我很多点拨和帮助，并且她相信我的能力让我去担任助教，每次授课时导师总会留给我一些时间，负责课程相关方面的内容，我觉得很荣幸。因为这一份信任，我很感谢她。现在我们不仅是师生，更是朋友。”而她的导师肖维青教授对她的研究课题、教学实践以及在博士生团队中的作用都非常欣赏，她说：“董琳娜把欧洲影视翻译研究最先进的理念带给了我们中国本土研究者，我们一起研讨，一起学习，每周线上线下研修活动对我的其他博士生都是一种启迪和推动；而且她的本科生助教工作卓有成效，广受学生欢迎，连我教了十年AVT的老法师，也经常感佩她在课堂中的教学活动设计。”



提及做助教的收获和对学弟学妹的建议，董琳娜谈到，“向学生教授知识的过程也是我提升自己的过程。一开始我讲课的时候，他们很感兴趣，但或许因为有点害羞，没有很多人愿意开口。后来我意识到这个问题，并试着调动他们的积极性、引导他们发言，慢慢的有一些学生开始主动分享一些他们的想法。这是一个很重要的点，因为我们老师应该做的就是让学生主动考虑事情，不只是我们一直传授知识，主动权也应该在学生手中。我觉得我们要让学生多想一想，在课堂上多发出自己的声音。”

从给班级带来fresh air 的西班牙学姐，变成后来调动班级氛围、传道授业解惑的助教，董琳娜收获颇丰；从初次接触崭新的研究领域，到勇敢探索，为爱发声，董琳娜坚定着公益的步伐。

未来可期

从学生，到助教，到青年榜样；从西班牙，到中国，再走向世界。或许我们的力量很小，可无数萤火之光聚在一起，也可与明月之辉相较。或许我们只是一支小小的蜡烛，但我们要以良知为烛台，知识为灯芯，思想就是我们的火焰，语言就是我们的光芒。文有道，字有神，语有力，言有灵，每一个外语人，都该像董琳娜一样，每一个人都可以用平凡的力量，发出非凡的声响。

采访、撰稿 | 盛晨 涂丹
排版 | 张丽娜



Views 237

Wow 1

Comment