Investigating the Multimodal Construal and Reception of Irony in Film Translation – An Experimental Approach

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Humanities

2017

Paulina Burczynska

School of Arts, Languages and Cultures

Table of Content

Word count: 74, 106

List of Tables

List of Figures

List of Acronyms and Abbreviations

SH1	*Sherlock Holmes*
SH2	*Sherlock Holmes: A Game of Shadows*
SUB_VO	Subtitles_Voice-Over
VO_SUB	Voice-Over_Subtitles
HLPs	High Level Participants
MLPs	Medium Level Participants
LLPs	Low Level Participants
AVT	Audiovisual Translation
AV	Audiovisual
TL	Target Language
SL	Source Language
TT	Target Text
ST	Source Text

Abstract

In the light of recent changes on the audiovisual scene in Poland, audiences can choose among different AVT modalities. Although voice-over still prevails on Polish TV, subtitles have become more and more popular as an alternative form of film translation on television. Due to rapid technological advances, commercial requirements and differences in Polish viewers' preferences, it is thus crucial to understand how audiences at different levels of English proficiency (low, medium, high) retrieve meaning, especially complex ironic meaning relayed through different methods of film translation, such as subtitles and voice-over and the extent to which verbal and non-verbal semiotic channels contribute to irony comprehension.

Wilson and Sperber's (1981, 1992; 1995) echoic theory of irony has been selected as the theoretical framework, given its ability to account for multimodal irony in audiovisual texts as well as the significant importance of non-verbal semiotic resources in the generation and interpretation of irony.

The study employs triangulation, incorporating descriptive, experimental and interactionist components. The descriptive component involves multimodal transcription (Baldry and Thibault, 2006) of selected fragments in which irony plays a pivotal narrative role. This procedure aims to determine what non-verbal modes contribute to the multimodal construal of irony and how it is relayed in the subtitled and voiced-over translations. In the experimental component, viewers' eye movements are recorded using eye-tracking technology while watching subtitled and voiced-over fragments of *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011). In the interactionist components, a questionnaire is used in order to elucidate how and/or whether they retrieve ironic meaning as intended by the filmmakers in the selected excerpts.

The most obvious finding to emerge from the descriptive data analysis is that multimodal irony is not relayed by the film dialogue alone but, rather, in unison with non-verbal semiotic resources. The instances of multimodal irony in the two Sherlock Holmes films were found to perform narrative and comedic functions by combining the visual, kinesic and acoustic modes of film language. The analysis and comparison of SL dialogues and TL translations revealed two broad categories of irony relay, namely: *preservation* and *modification*. The majority of the instances of multimodal irony were modified in the subtitled version, while preservation is only sporadically opted for. In its voiced-over counterpart, the intended meaning is preserved and modified in equal proportions. The experimental component showed major differences in gaze patterns among the participants with different language skills in the subtitled clips. For instance, on average, LLPs spent more time reading the subtitles than HLPs or MLPs. Similar visual behavior, on the other hand, was observed among all viewers in the voiced-over clips in which the on-screen character's face attracted the greatest amount of visual attention. The interactionist strand showed that the viewers retrieved the intended meaning to various extents depending on their English language proficiency. This data undergirds an assessment of the effectiveness of subtitles and voice-over in the translation and reception of multimodal irony on screen.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or institute of learning.

Copyright Statement

Acknowledgements

I would like to thank a number of people who supported me in the completion of this thesis:

- My Supervisor Dr. Pérez-González for his valuable feedback, undeniable support and advice during all stages of this research project. Thank you very much, Luis.

- All members of the AVT Lab at the Institute of Applied Linguistics, University of Warsaw, who kindly let me use their eye-tracking laboratory to conduct my experiment. I am grateful for all positive words, hospitality and guidance to make this study happened.

- My mum Anna and dad Ryszard whom I am truly thankful for encouraging me to continue with my studies although they have been missing me a lot.

- Several wonderful people like Tendai, Melanie, Frederick, Agnieszka, Peter, Wojtek, Daniel, Jackie, Isabelle and David and many others, who I met on conferences, workshops, courses and other academic-related events in the last years and who have become my friends. I could always rely on their piece of advice and support. I also would like to thank my dearest friends Magda, Paulina, Daga and Martyna who have always believed in me.

- My beloved Marcelo who has always been there for me when I needed, for his irresistible encouragement to finish this thesis and patience particularly during the last stages of writing up.

- Last but not least, I would like to thank all the participants who took part in my experiment, as without them it would have been impossible to complete this thesis.

Dedication

I would like to dedicate this thesis to my parents who have always encouraged me to pursue my dreams.

1 INTRODUCTION

In the contemporary world, digital technology is evolving at a frantic pace transforming the way information is created, processed and communicated. This technological revolution brought about the onset of digital terrestrial television and continued with "the expansion of video on demand, mobile and Internet TV" as well as HD and Blu-Ray DVD (Bogucki, 2010: 415). These advances have led to the development of new forms and features of translations in digital TV, for instance, subtitles for the deaf or audio description for the blind, offering audiences a possibility to adjust the reception of audiovisual texts to their needs or preferences (e.g., Pérez-González, 2014). More and more countries thus enable their spectators to choose among different types of audiovisual translation. As a result, viewers in dubbing or voice-over countries have turned to watch films with subtitles.

As target audiences with different profiles are constantly exposed to various forms of translation on screen, which have both advantages and disadvantages, it is highly important to examine how they process and understand the intended meaning in translated films. To date, a number of audience reception studies have been carried out incorporating psychological measures such as eye tracking[1] (e.g., Kruger and Steyn, 2014; Orrego-Carmona, 2014b; 2015; Szarkowska, et al., 2016) and EEG[2] (e.g., Kruger et al., 2013) to understand the influence of AVT modalities like subtitling on the cognitive processing and experience of audiovisual texts (Kruger, Szarkowska and Krejtz, 2015).

In the context of the audiovisual scene in Poland, audiences are exposed to voice-over on television and subtitles on DVDs, on the Internet and in cinemas. The present thesis is therefore an empirical study of how Poles with different levels of source language ability process and retrieve meaning, and more specifically irony, relayed via voice-over and subtitles, thereby it explores "audience processing of semantic information in various semiotic channels" (Gottlieb, 2005: 15). The current study also investigates the extent to which the verbal and non-verbal resources contribute to the participants' understanding of multimodal irony in the subtitled and voiced-over versions of action comedy genre, *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011), featuring a range of meaning-making resources across verbal, visual and acoustic channels. Hence, in addition to enhancing understanding of the impact of audiovisual translation modalities and linguistic

---

[1] The recording and study of the eye movements following a moving object, printed text, or other visual stimulus.

[2] Electroencephalography (EEG) is an electrophysiological monitoring method to record electrical activity of the brain.

abilities on comprehension of the complex phenomenon under study, this research projects aims to reveal which method of film translation is the most adequate to meet these viewers' needs.

While this research project examines the retrieval of multimodal irony using eye-tracking and questionnaires, and thereby concentrates on empirical and quantitative measures of audience response, it is important to realise that such reception always happens within a specific cultural context. Work in Cultural Studies, such as Hall (1997) and/or Pugliese (2010) has pointed to the part that existing cultural knowledge and experience plays in the construal of the intended meaning. However, one early finding of this research is that emotional response to irony tended to be very similar between viewers from Poland and those from the United States. More details of this finding and the methods used to generate this data can be found in Chapter 7.

Similarly, to other European countries, Poland has also long audiovisual traditions that determined the current state of affairs in the audiovisual market. In order to contextualise the research questions in the present study, this thesis will now turn to an outline of the history of AVT in Poland placing particular emphasis on the changes that have been made in the Polish audiovisual scene in the recent years.

1.1 History of AVT in Poland

Polish audiovisual traditions date back to pre-war Poland, when dubbing used to be employed to translate foreign language films. In 1980s, during the communist era, voice-over became the prevailing form of audiovisual translation, predominantly for economic reasons (Szarkowska, 2009; Bogucki, 2010; Sepielak, 2016). Viewers of voiced-over content can still listen to the original soundtrack for one or two seconds before the narrator's voice is superimposed on the original dialogue.

During the communist era, Poland's access to the Western audiovisual industry was extremely limited. In the 1976 - 1989 censorship period, 60% of films were produced in other communist countries, while foreign films were submitted to the verification process in which a censor could approve or reject a film due to "unwanted ideological content" producing bowdlerised versions of films (Misiak, 2006; Szarkowska, 2011: 161). Certain films could be banned in their entirety, specific scenes or sounds could be cut out or drowned out (Misiak, 2006). Although Poland is no longer under a communist regime, some forms of social censorship in relation to the depiction of sexuality and the use of offensive language on TV have remained. Consequently, audiovisual translators are obliged to abide by these

regulations in their translations (Garcarz, 2007; Szarkowska, 2011). Although communism was brought down with the fall of the Berlin Wall on 9th November 1989, "the post-soviet voice-over translation of fiction still remains uncharted territory" in this part of Europe (Hołobut, 2015: 225).

While in Western European countries voice-over is used to render factual television programs like science or history documentary or interviews, in Poland as well as other Eastern European countries (such as Russia, Ukraine, Latvia, Lithuania, and Estonia) voice-over tends to be used to translate fiction films on television. Apart from the historic and economic factors that determined the use of voice-over on Polish TV, there are also sociological aspects that hinder a wider application of subtitles on television in Poland.

The use of a single male voice artist (called *lektor* in Polish) to produce the voice-over narration known as "szeptanka" (BT: "whispered interpreting") among audiovisual translators is still favoured by and deeply rooted in the mentality of Polish society (Szarkowska, 2009). While a monotonous voice simply 'reading' the dialogues over the original soundtrack does not sound 'strange' at all for Polish viewers, critics believe that in voice-over "the characters in the film lose their identity and acting quality can only be transmitted visually and not orally" (Dries, 1995: 6 in Szarkowska, 2009: 186), despite the access to "some of the original flavour" of the film (Gottlieb, 2005: 25). As a result, the narrator's voice may hinder the reception of the characters' accent, intonation, pitch or volume of speech as well as music and special effects that also significantly contribute to the construal of meaning on screen (Dix, 2008; Philips 2000).

The fact that the gender of the voice artist chosen to narrate a film might not necessarily be the same as that of the on-screen character being translated has also attracted criticisms. While in certain Eastern European countries (e.g., Lithuania, Russia, Ukraine) a female voice artist will be relied upon to read the dialogues lines delivered by female on-screen characters (Sepielak, 2016), the practice in Poland depends on the genre being translated. When translating non-fiction productions, the *lektor* can be either male or female, for instance, female voices are preferred in cooking or nature programmes (Szarkowska, 2009). However, "regardless of the gender of the screen character, the dialogue in fiction films in Poland will always be read out by a man" (Szarkowska, 2009). Although there is no concrete reason underlying this choice, male voice artists may have been used due to solely aesthetic reasons (Sepielak, 2016) as women's voices are considered to convey more subjective feelings that were deemed unacceptable in voice-over narration (Rodkiewicz-Gronowska see in Kotelecka, 2006). Additionally, when viewers are exposed to the *lektor's*

unusual timbre or pitch voice, their experience of watching voiced-over films can be diminished (Woźniak, 2012).

The retrieval of meaning may become even more difficult when the dialogue involves more than two on-screen characters within a given scene (Tomaszkiewicz, 2006), and hence several characters may be speaking at the same time. Since it is not feasible for the *lektor* to read out everything, "the target text inevitably undergoes reduction and condensation" (Szarkowska, 2009: 186). The most commonly omitted elements from the original soundtrack range from vocative forms of address, greetings, borrowings like "OK" to swearwords and/or repetitions (Tomaszkiewicz, 2006: 118-119). The fact that the *lektor* reads only part of the original dialogue can be spotted by viewers with a high degree of source language proficiency (Mirecka, 2002), and can even hinder or mislead their comprehension of the audiovisual text.

On account of such technical and translational drawbacks, voice-over has been regarded as "a technique without a future" (Tomaszkiewicz, 2006: 102) and "inferior to the two other types, subtitling and dubbing" (Bogucki, 2010: 415) – its only advantage being that it is cheaper than other more mainstream forms of audiovisual translation (e.g., Garcarz, 2007). Nevertheless, the vast majority of Poles has supported voice-over translation on TV in Poland for many years. A study conducted by the Institute SMG/KRC Poland[3] in 2002 confirmed that a great number of responders favoured voice-over (50.2%) over dubbing (43.3%) and subtitling (8.1%) (Bogucki, 2004). The reason for this aversion to subtitles turned out to be very simple: they read subtitles very slowly at the time (Mirecka, 2002), as the historical prevalence of voice-over prevented them from developing these essential fast-reading skills leading to a situation where "Polish viewers are not ready to accept subtitles in everyday practice" (Garcarz, 2007: 131). A few years later, in 2007, another survey commissioned by the public television (TVP) demonstrated that voice-over was still the preferred modality of audiovisual translation on TV, while subtitles were still supported by only 4% of those surveyed (Garcarz, 2007).

1.2 Changes in the Polish Audiovisual Landscape

The year 2008 witnessed considerable changes on the audiovisual scene in Poland. The Polish media began to lobby for the inclusion of open interlingual subtitles[4] on television,

---

[3] SMG/KRC Institute of Market Research and Public Opinion Research is the largest marketing research centre in Poland and throughout Central and Eastern Europe.
[4] Open subtitles are always displayed on the screen and cannot be removed, even when requested (Debevere, *et al.* 2011).

drawing on the educational benefits for foreign language learning while watching programmes, films and series in the original soundtrack with subtitles (e.g., Díaz Cintas and Fernández Cruz, 2008; Szarkowska, 2009). As a result, the authorities of public television channel, TVP2, decided to broadcast an American TV series for teenagers with subtitles for the first time in October 2008 (Szarkowska, 2009). As advocated by Wojciech Pawlak (director of TVP 2 in 2008) subtitling can be described in the following terms "po prostu znakomity sposób na naukę języka angielskiego" (BT: "just an excellent way to learn English") (Sowa, 2008). Following this trend, some private film channels such as *Ale Kino!* also enabled viewers to watch certain films in the original version with closed[5] interlingual teletext subtitles. These first attempts to introduce subtitles increased support for the inclusion of this type of audiovisual translation on television. In 2008 a survey conducted by TNS OBOP[6] revealed that subtitles were preferred by 19% of society, and interestingly, nearly one-third of those surveyed were younger than 29 years (Sowa, 2008).

This strong trend towards subtitling has been maintained in Poland "where a growing number of subtitled films can be enjoyed on private TV channels and in cinemas" in the following years (Chaume, 2013: 116). It was therefore hoped that the launch of digital terrestrial TV would accelerate the introduction of subtitles on a larger scale and provide viewers with a choice on whether to watch TV programmes with voice-over or subtitles. Contrary to the expectations of several Polish audiovisual translation scholars (e.g., Bogucki, 2010; Szarkowska, 2009), the launch of digital television channels on 23rd July 2013 has not revolutionised the audiovisual market significantly. That is to say, voice-over still reigns on TV and viewers are still unable to choose their preferred audiovisual translation modality on television.

Surprisingly, the untapped potential of digital television was spotted by one of the Polish political parties. As *Dziennik Gazeta Prawna*[7] (2013) reported, PSL[8] put forward a Bill seeking to impose on both public and private television stations the obligation to provide closed captioning, thus boosting the acquisition of foreign languages while watching films. Despite these attempts, this Bill was not passed in parliament mainly for economic reasons. Once it transpired that the government would not make a financial contribution to deploy

---

[5] Closed subtitles are displayed on the screen optionally and can be removed, when requested (e.g., Teletext) (Debevere, *et al.* 2011)

[6] TNS OBOP is a market research agency examining public opinion in Poland.

[7] *Dziennik Gazeta Prawna* (meaning *Daily Legal Newspaper* in English) is a Polish newspaper headquartered in Warsaw and published in Poland from Monday to Saturday. The paper focuses on economic and legal affairs.

[8] *Polskie Stronnictwo Ludowe*, abbreviated to PSL (traditionally translated as Polish Peoples Party) is a political party in Poland focused on development of agriculture and social-market economy.

the technological platforms required to enable viewers to choose between subtitling and voice-over, both public and private television stations refused to cover these costs (*Dziennik Gazeta Prawna,* 2013). In parallel to this, public and private station executives (*Dziennik Gazeta Prawna,* 2013) remained convinced that audiences still favour voice-over over subtitling.

Along with growing language awareness, Poles have started to declare themselves in favour of broadcasting more subtitled programmes on television channels in Poland (Szarkowska and Laskowska, 2014). This growing shift towards subtitles has been particularly observed among well-educated younger popultion in their 20s and 30s, who opt for subtitles similarly to their Spanish and Italian counterparts (Chaume, 2013). A study carried out in 2014 as part of the HBBTV4AL[9] project demonstrated that 77.25% of those surveyed (aged 21-40) opted for subtitles as their favoured modality of audiovisual translation on TV, while only 6.88% expressed a preference for voice-over (Szarkowska and Laskowska, 2014).

One of the main reasons behind the evolution of viewers' tastes in recent years is a gradual increase in the level of foreign language skills. English has become a widely taught and studied foreign language in Polish schools and universities since the fall of the Berlin wall (Karoń, 2006). As a result, a growing level of English proficiency has been observed, in particular among students and graduates, in their 20s and 30s, as the report published by the Opinion Poll Agency[10] in 2013 revealed. The growing percentage of Poles who are familiar with English also underpins the choice of subtitles over voice-over.

To date, there have been several empirical investigations into the reception of interlingual and intralingual subtitles among Polish hearing, hard of hearing and deaf viewers that brought a wealth of highly important scientific results (e.g., Szarkowska et al., 2016). Nevertheless, no previous study has investigated and compared how hearing audiences with a high, medium and low level of English interact with the multimodal texts like films to retrieve meaning, in particular more complex meaning like irony relayed with subtitles or voice-over. Understanding how participants with different linguistic skills process the subtitled and voiced-over films will allow to identify the type of audiovisual translation that

---

[9] The HBB4ALL (Hybrid Broadcast Broadband for All) is a European project co-funded by the European Commission under the Competitiveness and Innovation Framework Program (CIP) by 12 partners. They are from several complementary fields: universities, TV channels/broadcasters, research institutes, and SMEs. All are experts in the field of media accessibility and the multi-device environment. The project started in December 2013 and runs for 36 months.

[10] The most recent study regarding the levels of English proficiency among Polish students and university graduates was conducted by Opinion Poll Agency and commissioned by the authorities of the city of Szczecin in Poland and it is available at: http://bip.um.szczecin.pl/UMSzczecinFiles/file/WOIiB/Broszura_PL.pdf

is the most optimal for their abilities. These unexplored issues will be addressed through the research questions presented in the section below.

1.3 Research Questions

In the light of the current situation in the Polish audiovisual scene, the needs of viewers with different linguistic abilities are addressed in the research questions, which aim to identify the type of AVT that supports best different audience profiles. To achieve these research goals, the present study examines the construal, relay and cross-cultural reception of instances of multimodal irony identified in Sherlock Holmes films under analysis. These three foci are expressed in the following set of secondary questions in order to answer the central research question:

Central research question:

What is the contribution of verbal and non-verbal resources in irony comprehension by Polish viewers in the subtitled and voiced-over Polish versions of *Sherlock Holmes* (2009) *and Sherlock Holmes: A Game of Shadows* (2011)?

Secondary research questions:

1. How is irony construed in the two Sherlock Holmes films?
   a) How do verbal resources contribute to the construal of irony?
   b) How do non-verbal resources contribute to the construal of irony?

2. How is irony relayed in the Polish subtitled and voiced-over version of the two Sherlock Holmes films?

3. How do Polish viewers, representing different levels of English proficiency, consume the intended ironic meaning in the subtitled and voiced-over Sherlock Holmes films, as shown by eye-movement data?

4. To what extent are Polish viewers, representing different levels of English proficiency, able to grasp the intended ironic meaning in the subtitled and voiced-over Sherlock Holmes films, as shown by questionnaire responses?

5.   What type of film translation, subtitling or voice-over, proves to be optimal for the reception of multimodal irony by audiences representing different levels of English proficiency?

Responses to **Questions 1-2** will be provided by the analysis of descriptive data. The data set consists of 6 instances of multimodal irony identified in the films *Sherlock Holmes (2009)* and *Sherlock Holmes: A Game of Shadows (2011).*

Question 1 explores the way multimodal irony is construed in the two Sherlock Holmes films. Given that a film is a complex polysemiotic medium (Chaume, 1997), film dialogue is not capable of construing meaning in isolation from the non-verbal modes realised through the language of film (Kozloff, 2000). Additionally, the composition of irony is also believed to be supported by a number of non-verbal kinesic and acoustic markers (e.g., Rockwell 2000; Attardo et al. 2003; Gibbs 2011). It is anticipated that ironic meaning will be produced through the amalgamation of both the verbal and non-verbal modes in the film text under study. The first subquestion (a) refers solely to the contribution of the verbal mode to the construal of irony through film narration and its functions. Since irony can be considered a manifestation of "linguistic indirectness" (Mernit, 2001), it is important to identify the linguistic tools and narrative functions the filmmakers employ to convey the intended ironic meaning on screen. The second subquestion (b) is concerned with the way multimodal irony is construed through the co-deployment of non-verbal modes of film language pertaining to mise-en-scène, cinematography, editing and sound. Here, Sperber and Wilson's echoic theory of irony is applied for the analysis of how non-verbal elements typically express irony in the film text. In order to examine how the verbal and non-verbal modes intersect to convey the intended ironic meaning in a multimodal way in the clips under analysis, I will use an adapted model of multimodal concordance (Baldry and Thibault, 2006), as described in Chapter 4, to conduct a dissection of the elements that construe irony on screen and I will present the results from the descriptive data analysis in Chapter 5.

Question 2 investigates the transfer of multimodal irony in the Polish subtitled and voiced-over versions of the two Sherlock Holmes films. It explores the extent to which the translators either preserved or modified SL utterances when transferring the instances of multimodal irony from the original dialogues and debates whether the ironic effect has been retained (or not) in the TL utterances for the target audience. The transfer of ironic content has not yet been analysed in the voiced-over translation. In the subtitled version, however, modification (involving reduction, omission or paraphrase of the SL in the TL) has been

identified as the most common way of irony relay (Pelsmaekers and Van Besien, 2002), which involves a certain degree of manipulation of ironic meaning. The present project is intended to provide evidence as to whether modification is also observable in the subtitles or in the voice-over when multimodal irony is found in the original dialogue. The research hypothesis is that the audiovisual translators will tend to modify rather than preserve irony in the subtitled version and in its voiced-over counterpart, mainly due to the linguistic and cultural complexity of dealing with irony between the source and target languages and cultures, the multimodal nature of the films (meaning can be conveyed visually and/or acoustically) and to the modality of film translation (e.g., spatio-temporal restrictions). The ST and TT will be juxtaposed and analysed in tandem with the non-verbal elements in the multimodal transcription (Baldry and Thibault, 2006) and presented in the analysis of descriptive data set in Chapter 5.

Experimental sessions are carried out in order to answer **Questions 3**, **4** and **6**. In the present thesis, the inclusion of the experimental and interactionist components in the study of multimodal irony allows to verify the validity of the multimodal analysis of the descriptive data by gauging the extent to which multimodal irony, identified in the two Sherlock Holmes films, is recovered by Polish viewers. Additionally, this project compares the recovery of the intended meaning with subtitles and voice-over by the participants at three different levels of English proficiency (high, medium, and low).

Question 3 investigates how the subjects distribute their visual attention across the screen when watching the clips featuring the use of irony with subtitles and voice-over. The analysis of their viewing behaviour is performed via eye-tracking technology which is described in Chapter 4 in greater detail. Thus, Question 3 is intended to identify some characteristics of eye movements that are related to irony comprehension, which is either preserved or modified in the subtitled or voiced-over clips. This data is also expected to foreground differences in gaze patterns between the participants with a high, medium and low level of English in the perception of verbal and non-verbal visual elements in the subtitled and in voiced-over films and to test the hypotheses regarding the distribution of visual attention in the clips with subtitles and voice-over. The results of the eye-tracking analysis are provided in Chapter 6.

Question 4 examines the degree to which the participants retrieve ironic meaning as intended by the filmmakers in the two films under scrutiny. Concurrently, it will reveal whether the analyst's expectations prove to be correct as to how well Polish viewers with three different levels of English proficiency comprehend irony and the extent to which the

non-verbal cues contributed to recovery of irony relayed with subtitles and voice-over. At the same time, the hypothesis regarding the comprehension of multimodal irony will be tested. The participants' responses were collected with a questionnaire especially designed for the purpose of this experiment (see Chapter 4) and analysed in Chapter 7.

Question 5 seeks to determine which of the audiovisual translation modalities used to relay multimodal irony in the current project is more effective for allowing audiences with a given profile to retrieve the intended ironic meaning. Drawing on the results obtained from the experimental and interactionist component, the answer to Question 5 will be discussed in Chapter 8.

Finally, the **Central Research Question** aims to ascertain in Chapter 8 the extent to which the verbal and non-verbal semiotic resources which have been found to contribute to the construal of multimodal irony are retrieved by Poles in the subtitled and voiced-over versions of Sherlock Holmes films. It is assumed that information conveyed visually and acoustically can facilitate comprehension of the intended ironic meaning on screen.

1.4 Structure of the Thesis

The present the thesis consists of seven chapters:

Chapter 2 examines three influential pragmatic approaches to the process of production and reception of verbal irony in order to select the theory that can guide the present study of multimodal irony in the film text the most effectively. It starts with the conceptualisation of verbal irony from the perspective of Grice's theory of conversation (1967/89) and continues with the enhancement of his account put forward by Clark and Gerrig (1984) in the context of the Pretense Theory. Then, the focus is shifted from philosophical pragmatics to the cognitive pragmatics and discusses the Echoic Theory of Irony proposed by Sperber and Wilson within Relevance Theory (Sperber and Wilson, 1992, 1995, 2012). The chapter proceeds to explain in detail how the central concepts developed within Sperber and Wilson's framework support the relay of ironic meaning with non-verbal elements. Their theory of irony will constitute the basis for the investigation of multimodal irony in the subtitled and voiced-over films within Audiovisual Translation Studies. The chapter ends with a discussion of the capabilities of each of the three theorisations in the process of generation and comprehension of ironic meaning in the context of the current project.

Chapter 3 explores how the verbal and non-verbal elements intersect to construe ironic meaning on screen. First, it outlines the contribution of non-verbal markers of irony within the field of psychology, and then moves on to its conceptualisation from the perspective of cultural studies. The chapter further examines how non-verbal elements work in unison drawing on Baldry and Thibault's (2006) approach to multimodality. After delineating the tenets of the multimodal account, the next section discusses film studies literature in order to see how the ironic intent can be expressed by a combination of elements pertaining to the language of film, i.e., mise-en-scène, cinematography, editing and sound. Although, particular emphasis is placed upon the non-verbal cues, a separate section is devoted to film narration and its functions in the construal of multimodal irony in the film text. The section that follows is concerned with the role of *pastiche* and intertextuality in irony relay in the genre of action comedy represented by the film text under scrutiny. The chapter concludes with a proposed definition of multimodal irony in the polysemiotic context of film drawing on insights from Sperber and Wilson's operationalisation of irony.

Chapter 4 describes the design and implementation of the methodological apparatus for the study of multimodal irony in the subtitled and voiced-over film text proposed in the present thesis using a triangulation of methods. Before describing the methodology in detail, the criteria for selecting participants for the empirical study are discussed and the rationale underpinning the decision to choose *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011) as descriptive data is outlined. Then, the chapter proceeds to elaborate on each stage of the methodological framework, i.e., multimodal transcription, questionnaire and eye-tracking. Furthermore, the chapter illustrates how a set of methodological tools (including multimodal transcription and questionnaire design) is deployed on the basis of the examples selected from the current data set. The final section of the chapter provides a comprehensive account of how the eye-tracking experiment was planned and carried out followed by an explanation of experimental data processing.

Chapter 5 is dedicated to the exploration of descriptive data and seeks to address the first set of research questions presented above. It embarks upon a detailed analysis of the verbal and non-verbal elements of the language of film that concurrently relay

the intended ironic meaning in a multimodal way. The second part of the chapter explores how multimodal irony is relayed in the subtitled and voiced-over versions of two films by examining and comparing the English dialogue with subtitles and voice-over. The identified categories of irony relay are analysed and illustrated using examples drawn from the descriptive data set. The last part of the chapter interprets the findings on irony relay in the light of considerations related to the type of translation and the level of English proficiency of the participants.

Chapter 6 is concerned with the examination of eye-tracking data in response to the second set of research questions outlined above and aims to provide a verification of the findings from descriptive data. In the first part of the chapter, eye-movement data from the participants are examined and interpreted for each of the clips with subtitles and voice-over in which multimodal irony was predominantly conveyed visually. The second part of the chapter explores the participants' viewing behaviour in the subtitled version and in its voiced-over counterpart in which multimodal irony is principally transferred acoustically. The chapter concludes with a discussion of the most important findings in light of the situation of the Polish audiovisual landscape.

Chapter 7 presents a qualitative and quantitative analysis of participants' responses to the questionnaire and aims to provide answers to the third set of research questions. This chapter examines the extent to which the viewers were able to retrieve meaning, and more importantly ironic meaning, and provides an account of the non-verbal elements contributing to their comprehension. Thus, the first and the second sections outline similarities and differences in the understanding of multimodal irony relayed visually and acoustically, respectively, with subtitles and voice-over by means of pertinent examples. At the same time, the contribution of non-verbal elements such as facial expressions, film score or camera position to irony recovery is considered in regard to the type of translation and the level of English proficiency. The chapter continues with a discussion of nonverbal elements construing irony that may be deemed culturally specific and concludes with a summary of findings.

The concluding chapter of the thesis, **Chapter 8** summarises the main findings and provides responses to the central question of this research project as well as the subquestions mentioned above. Additionally, it refers to some of potential implications and elaborates on the contribution of the thesis to the current fields of study. Finally, the chapter also suggests some future research avenues that can be explored further.

2 PRAGMATICS OF IRONY

2.1 Introduction

Irony as a linguistic phenomenon has been investigated for over 2,500 years. Throughout this period, scholars of philosophy, rhetoric, literary and pragmatic-linguistics studies have attempted to conceptualise, classify and define the notion of irony. Although, earlier scholars contributed significantly to the study of irony, pragmaticists (e.g., Grice, 1974; Clark and Gerrig, 1984; Wilson and Sperber, 1981, 1992; 1995) specifically explored how speakers produce ironic utterances and how listeners grasp the intended ironic meaning, particularly when the latter is conveyed implicitly. In addition, several pragmatic scholars have also attempted to gauge to what extent non-verbal elements support the relay of ironic meaning in the context of conversational interactions (Grundy, 2000).

The present analysis of multimodal irony is guided by the conceptualisations of irony put forward by scholars in the field of pragmatics as these provide a theoretical basis for understanding how ironic meaning is construed on screen and received by Polish audiences. It is necessary, however, to evaluate which pragmatic-driven approach guides the analysis of multimodal irony the most effectively in the context of this study.

This chapter therefore sets out to look at three main pragmatic accounts of irony in order to examine and compare their ability to explain the processes behind its production and interpretation. The purpose of this overview of alternative approaches is to provide the basis for a sound and informed decision as to the theory of irony used in this study to account for a wide range of instances of irony, while also including non-verbal elements in the context of the analysis of multimodal irony in the film text. Section 2.2 examines the notion of verbal irony from the perspective of Grice's theory of conversation (1967/89), while Section 2.3 interrogates the account of verbal irony proposed by Clark and Gerrig (1984) in the context of Pretense Theory. Section 2.4 examines Sperber and Wilson's Echoic Theory of Irony (1981; 1992; 1995) developed within Relevance Theory (1986; 1992). This chapter ends with Section 2.5, which analyses the relay of ironic meaning in audiovisual productions within the domain of AVT highlighting the role that non-verbal resources play in the transfer of irony in subtitled film texts.

2.2 Grice on Verbal Irony

This section describes the phenomenon of verbal irony in the view of Grice's theory of conversation (1967/89) in order to explain the conditions in which utterances are deemed to

be ironic, and thereby to examine the extent to which the theoretical framework put forward by Grice can be applied to guide the present analysis of multimodal irony.

2.2.1 Construal and comprehension of irony

Although Grice's account of irony still resembles the standard definition of irony put forward by ancient philosophers, according to Kaufer (1981: 499), Grice's theory constituted a "significant advance" in an explanatory basis of how irony is generated and interpreted within the domain of pragmatics.

In pragmatics, it is argued that successful and effective communication with one another involves a common aim of conversation and a certain degree of cooperation between participants particularly when meaning is communicated indirectly like irony (Daly, 2013). In his *Logic and Conversation,* Grice (1975) proposed an overarching principle known as the Cooperative Principle (CP) to describe cooperation between speakers and listeners in conversational interaction. The Cooperative Principle is thereby viewed as a keystone of conversation that directs participants to construct rational, productive and meaningful communication.

Under the umbrella of the Cooperative Principle, Grice (1975: 45) proposed a set of more specific norms, his "Maxims of Conversation", which attempt to describe how speakers behave in conversation. These maxims are: Quantity, Quality, Relation and Manner. These maxims often consist of several conditions. According to the Cooperative Principle, at least one condition needs to be fulfilled by participants in conversational interactions in order to follow the rational flow of conversation. It is necessary here to define each of the maxims and present how they help govern the conversations plausibly. For example, the maxim of Quantity (i.e., Make your contribution as informative as is required) outlines that participants should not provide too much or too little information as some extra or incomplete information may lead listeners to an incorrect interpretation of an utterance. The maxim of Quantity has been illustrated in Example 2.1:

Example 2.1

> If Mary asks Peter to lend her £10, she expects him to give her exactly £10 rather than £5 or
> £20

(Daly, 2013)

The maxim of Quality (i.e., Do not say what you believe to be false) says that participants have to produce true utterances. Alternatively, they are expected to adduce evidence to support their utterance. Otherwise, their communication or behaviour is considered as uncooperative. The maxim of quality has been briefly presented in Example 2.2:

Example 2.2

> If Mary asks Peter for a cookbook while preparing dinner, she does not expect him to give
> her a novel.
>
> (Daly, 2013)

The maxim of Relation (i.e., Be relevant) indicates that participants are expected to produce appropriate and relevant utterances to the given situation. The illustration of a situation in which B's response is not relevant to A's utterance has been given in Example 2.3:

Example 2.3

> A: I saw an amazing programme on BBC3 last night.
> B: I've got a really weird itchy feeling behind my ear.
>
> (Clark, 2013)

In Example 2.3, B's response could trigger all sorts of interpretations. For instance, B could indirectly communicate that he/she is not interested in the new programme that A watched last night or indicate that he/she prefers to change the topic of conversation. The maxim of Manner (Avoid obscurity of expression) generally informs how participants should formulate their utterances and not what they should express in an utterance. Clark (2013) exemplifies it by saying that it would be strange if someone is simply asked if he or she is tired and instead of a clear, brief answer 'yes', gave an elaborative and complicated statement like in Example 2.4:

Example 2.4

> I am uttering my belief and reporting my own sensation when I tell you in answer to your
> question that tiredness is something which I am currently experiencing.

Grice (1967/89) argues that participants should follow the above maxims to have a rational and successful conversation with one another and to avoid misunderstandings. The following Example 2.5 illustrates a conversation in which participants are abiding by all 4 maxims and thus manage to communicate successfully:

Example 2.5

> A: Where are the car keys?
> B: They're on the table in the hall.
>
> (Thomas, 1995)

In Example 2.5, the speaker has responded unambiguously (Manner), honestly (Quality) conveying the exact amount of information (Quantity) and giving a relevant response to A's question (Relation). Since the speaker B has accurately and directly communicated what was meant, the listener A does not have to look for another interpretation.

In the context of Grice's framework (1975), the flouting of a maxim is a highly significant condition in the generation of implied meaning. Grice (1975) argued that the

31

flouting of a maxim occurs when speakers blatantly fail to observe the maxim because they intend to provoke hearers to look for the implied meaning (what is meant) of an utterance which differs from the expressed meaning (what is said). The alternative meanings implied in such cases are called implicatures[11]. For example, instead of a clear direct response to a question, speakers could produce a totally unrelated utterance indirectly using metaphor, hyperbole, understatement or verbal irony, for example, communicating something else, and thereby generating different kinds of conversational implicature.

Ironic meaning is thus generated when speakers deliberately flout a particular maxim, namely, the Maxim of Quality in conversational interactions. One example might be to say "X is a fine friend" of someone being disloyal because he/she betrayed all business secrets to the competing company (Grice, 1967). In this case, the speaker failed to observe the Maxim of Quality by conveying a blatantly false statement, thus disobeying the convention of literal truthfulness. As a result of this flouting, speakers produce a related proposition that is contrary to what is literally said and for listeners, the indirect meaning to infer irony must supersede the literal meaning:

> [in irony] it is perfectly obvious to A [the speaker] and his audience that what A has said or has made as if to say is something he does not believe, and the audience knows that A knows that this is obvious to the audience. So, unless A's utterance is entirely pointless, A must be trying to get across some other proposition than the one he purports to be putting forward. This must be some obviously related proposition; the most obviously related proposition is the contradictory of the one he [the speaker] purports to be putting forward. (Grice, 1975: 53)

As mentioned above, Grice (1986) understood irony as a related proposition, which "is expressing the opposite of what the speaker actually intends", as shown in Example 2.6 singled out from the present data set in which Sherlock is mocking Inspector Lestrade:

Example 2.6

> Watson visits Sherlock after a longer absence. He calls his name but hears no response. Watson thus takes a sit on a chair and starts reading a newspaper. When he sits relaxed, Sherlock takes him by surprise and shots an arrow straight into Watson's newspaper. Once Sherlock finishes his performance and stops hiding, Watson comments without looking at him:
>
> **Watson:** Oh, how I've missed you, Holmes.

(*Sherlock Holmes: A Game of Shadows*, 2011)

---

[11] "Implicature" denotes either (i) the act of meaning or implying one thing by saying something else, or (ii) the object of that act. (Source: Stanford Encyclopedia of Philosophy)

In Example 2.6 Watson conveys a blatantly false proposition that is the contradictory of what he really intends to say indicating that he was not really missing Sherlock's extravagant behaviour.

According to Grice (1967/89) a two-stage process is essential in order to retrieve ironic meaning successfully. First, listeners have to identify that speakers are flouting of the maxim of truthfulness (Quality) specifically which is recognised as a signal for listeners to reject the literal meaning. Second, listeners have to deduce the intentions of the speakers who relay ironic meaning. Nevertheless, Grice's early definition of irony does not always work in practical analysis since in some instances the ironic effect does not seem to be achieved, as presented in Example 2.7:

Example 2.7

> A and B are walking down the street, and they both see a car with a shattered window. B says, *Look, that car has all its windows intact.* A is baffled. B says, *you didn't catch on; I was in an ironical way drawing your attention to the broken window.*
>
> (Grice, 1989: 53)

In Example 2.7, although the utterance produced by speaker B meets Grice's conditions for generating and understanding irony, i.e., speaker B makes a blatantly false statement and conveys the opposite of the literal meaning, irony still does not seem to arise. The reason for this may be linked to the absence of a certain type of attitude or judgment expressed by the speaker. Yet Grice limited the purpose of irony to the expression of negativity and criticism only acknowledging that an ironic remark should be contemptuous in nature as he highlighted in *Further Notes on Logic and Conversation:*

> irony is intimately connected with the expression of a feeling, attitude, or evaluation. I cannot say something ironically unless what I say is intended to reflect a hostile or derogatory judgment or a feeling such as indignation or contempt.
>
> (Grice, 1989: 53-54)

As a result, it can be complex for a listener to interpret the utterance "Look, that car has all its windows intact" (Example 2.6) as ironic since "it is hard to see what could have justified this critical judgment or attitude in the circumstances described" (Wilson 2006: 8). Thus, it is difficult to identify a type of critical attitude that the speaker B was expected to express towards the speaker A.

Grice (1989) also noticed that the expression of irony tends to involve the use of a specific tone of voice that should be usually contemptuous in nature. Yet he was very inconsistent in the theoretical considerations on the ironic tone of voice. On the one hand, Grice deemed it was obligatory underlining that "if speaking ironically has to be, or at least appear to be, the expression of a certain feeling or attitude, then a tone suitable to such a

feeling or attitude seems to be mandatory" (Grice, 1989: 54). On the other, he doubted that a tone of voice exists as a separate entity. Grice did not elaborate on the tone of voice underpinning the retrieval of ironic meaning any further yet suggested that a series of experimental and empirical studies should be conducted to see if this tone of voice qualifies to be labelled as ironic.

2.2.2 Objections to Grice's study of irony

Some theorists, such as Clark (2013) view Grice's contributions to the theorisation of irony within the domain of pragmatics to be very important. Others have argued that his contribution stimulated the subsequent discussion on the nature of verbal irony (Barbe, 1995) identifying several problem areas of the Gricean approach to the production and interpretation of ironic meaning within his account. This section will discuss the criticisms of Grice's theory that are most relevant to this present study of multimodal irony.

First, Grice's study of irony was frequently criticised for being inadequate (e.g., Kaufer 1981; Wilson and Sperber, 1981; Attardo, 2000; Wilson, 2006; Sperber and Wilson, 2012; Clark, 2013) particularly in the way ironic meaning is generated by speakers and interpreted by listeners. For instance, several scholars disagreed with Grice's account of irony (Kaufer, 1981; Wilson and Sperber, 1981; Utsumi, 2000) by arguing that flouting the Maxim of Quality is not the only vital condition necessary to generate ironic utterances. There are, indeed, several instances of irony in which ironic meaning is triggered by flouting Grice's other maxims or no maxim at all. This is supported by Sperber and Wilson (1981) who noticed if a speaker is not saying the opposite of what he/she means, the Maxim of Quality is not flouted but ironic effect still seems to arise. For instance, in some cases ironic meaning is generated when a speaker is not saying the opposite of what he/she intends to say but is saying "merely less than what is meant" construing irony through understatements (Sperber and Wilson, 2007: 36). In Example 2.8 the instance of irony has been shown that is construed with the use of understatement:

Example 2.8

The villain Lord Blackwood is sentenced to death, executed and buried in the cemetery. After a couple of days, however, Lord Blackwood disappears from the grave and Inspector Lestrade conducts an investigation to explain this unusual situation. When police officers open the coffin and it turns out that a different person was buried in Lord Blackwood's place, Inspector Lestrade says:

Inspector Lestrade: That's not Blackwood!
Sherlock Holmes: Well, now we have a firm grasp of the obvious.

(*Sherlock Holmes,* 2009)

In Example 2.8 Sherlock is saying that Lestrade has "a firm grasp of the obvious", Holmes is mocking him by tacitly saying that Lestrade is straining to even understand the obvious, not a good trait in a police inspector. Although Sherlock did not produce a blatantly false statement, the ironic effect still appears to arise. This example highlights the limitations of Grice's theorisation of irony, as the ironic effect is also achieved when the speaker produces understatements.

Second, some pragmatic scholars also questioned the truthfulness condition in the generation and interpretation of irony (Kaufer, 1981; Wilson and Sperber, 1981). That is to say, according to Grice (1967) all ironic utterances are considered to be blatantly false. In some instances, however, "ironists commonly speak the literal truth" (Kaufer, 1981: 499) and relay a true proposition in ironic utterances (Wilson and Sperber, 1981). Such an utterance can be thus understood as ironic, even though it relays a literally true meaning (Clark, 2013). As an illustration of the truthfulness condition, the following instance of irony has been presented in Example 2.9 from the present data set:

Example 2.9

Sherlock visits the villain Lord Blackwood in prison after Holmes caught him in the cathedral for the attempt of killing an innocent girl. While Sherlock wishes to extract more information from Lord Blackwood, he warns Holmes about the events that are still about to come.

Blackwood: I warned you Holmes, to accept that this was beyond your control, beyond what your rational mind can comprehend.
Sherlock: What a busy afterlife you're having.

(*Sherlock Holmes*, 2009)

In Example 2.9, Sherlock is mocking Lord Blackwood's statement with a literally true meaning that corresponds to a number of plans that Lord Blackwood has got in mind. Although the truthfulness condition was not violated, irony still seems to arise.

Similarly, these lines of reasoning can also be applied to the instances of irony when a speaker addresses the utterance to someone who did something unwise or other cases of self-irony in which speakers mock themselves. Although a speaker does not meet Grice's requirements, ironic meaning still seems to be triggered. Consider Example 2.10 chosen from the data set of the present study in which Sherlock is self-mocking his own unwise decision:

Example 2.10

Watson and Sherlock set out for a journey to find and kill the villain Professor Moriarty. On the way, they spend a night in the gypsy village who offer them horses so that they could reach their destination much faster. Since Sherlock cannot ride a horse, he chooses a little pony to ride. Because of his slow and little pony, Sherlock rides all the time at the end of the whole group self-mocking himself:

Sherlock: (riding a pony) Slow and steady wins the race.

35

Although Sherlock did not flout the Maxim of Quality or did not relay a blatantly false statement as shown in Example 2.10, the ironic effect still seems to arise through his self-mockery.

Given the examples provided above, it is clear that Grice's theory on the generation and interpretation of ironic meaning is very limited, as it only accounts for those instances of ironic content in which a speaker flouts the Maxim of Quality and says something that is blatantly false. The above example showed, however, that Grice's theory of irony is not able to account for these instances which cannot be explained with the help of the maxims, which is a serious weakness in respect to the current study.

Last but not least, Grice's theoretical approach does not acknowledge sufficiently the non-verbal resources that contribute to the construal and comprehension of ironic meaning. Admittedly, Grice did observe that a particular attitude or tone of voice may help to generate and interpret irony. However, he did not elaborate on these non-verbal elements any further and did not provide any examples to support his claim. Nevertheless, certain attitudes and tones of voice are critical to the present analysis of multimodal irony, as speakers may make listeners believe that the intended meaning is considered ironic. Still, Grice's thoughts provided a foundation for other theories of irony like the Pretense Theory put forward by Clark and Gerrig (1984).

2.3 Irony and Pretense

This section sets out to provide a critical analysis of Clark and Gerrig's (1984) approach to the study of irony encapsulated in the pretense theory, in which speakers are pretending to say something or to be a different person to relay irony, in order to see whether Clark and Gerrig's (1984) approach to the study of irony is capable of informing the present study of multimodal irony.

2.3.1 Clark and Gerrig's account of irony

The term "pretense" originally stems from the Greek word *eironeia*, which means "pretense" and "dissimulation" (Attardo, 2006). However, the operationalisation of irony as a form of pretense proposed by ancient Greek philosophers like Socrates or Plato differs significantly from other accounts of irony developed by succeeding scholars and philosophers.

In his work, Grice (1967/89) partially included Plato's idea of irony as pretense. In contrast to ancient Greek philosophers, Grice did not interpret pretense as form of deception, but rather as an instrument to convey ironic meaning. Nevertheless, he did not explain how irony is generated and interpreted through pretense. Clark and Gerrig (1984: 121) thus attempted to broaden Grice's account of irony as pretense answering the question: "What is the ironist pretending to do?" and illustrating the functions and mechanism of irony, i.e., how ironic utterances are generated by speakers, who produce ironic utterances, and understood by listeners, who receive ironic utterances through pretense.

The central idea behind pretense is that speakers of an ironic proposition are not themselves producing a speech act but they are pretending to perform it in order to relay a mocking or contemptuous attitude to the speech act itself (Clark and Gerrig, 1984). In other words, speakers are pretending to be a different person, whether real or imagined, through imitation of his/her characteristic gestures, tone of voice or content of speech, among others. Concurrently, speakers are dissociating themselves from the person who in fact they are not. According to this view of irony, speakers change their role into the role of pretenders to convey ironic meaning and consequently, abandon their own voice in exchange for the new ironic voice of the person they pretend to be (Barbe, 1995).

In Clark and Gerrig's view, pretense is the key condition to understand irony. Listeners have to recognise when speakers are pretending, what type of person they are pretending to be and whom they are pretending to address (Clark and Gerrig, 1984; Wilson, 2006; Kalbermatten, 2006). In this view, Clark and Gerrig (1984: 124) contend that "ironists can pretend to use the words of any person or type of person they wish, just as long as they can get the intended audience to recognise the pretense". Nevertheless, if speakers inform their listeners that they are making a "pretense", e.g., they announce whose tone of voice they are imitating, the ironic effect will be lost.

Pretense theoreticians further argue that "common ground", i.e. "mutual beliefs, mutual knowledge, and mutual suppositions", between speakers and listeners is the key factor to understand pretense and interpret irony correctly because "speakers are not just ironic: They are ironic only to certain listeners", particularly to those who share speakers' point of view or beliefs (Clark and Gerrig, 1984: 124). As a result, listeners' understanding of irony does not solely depend on their ability to realise that speakers are pretending to be someone else, as irony comprehension also heavily depends on the information and knowledge shared between speakers and listeners. Example 2.6 below illustrates a short satire in which the importance of the "common ground" between speakers and listeners is

37

essential to understand irony. In this example, the writer Craig Brown is pretending to be American President Barack Obama and imitates his speech while discussing cornflakes. In order to retrieve the ironic meaning behind it, listeners have to be familiar with the speech style that Obama is famous for, as Example 2.11 shows:

Example 2.11

> These cornflakes are real and they are everywhere. And I tell you this, Michelle, I say. The packet may have been shaken, but the flakes will recover. So, it is with profound gratitude and great humility that I accept my breakfast cornflakes.

(Craig Brown *The Lost Diaries.* 2010. Fourth Estate, London, cited in Clark, 2013: 288)

As presented in the example above, Clark and Gerring (1984) attempted to elaborate on the non-verbal dimension of irony, namely the ironic tone of voice, to support the generation and interpretation of irony through pretense. Clark and Gerring (1984) explained that when speakers turn into pretenders, they also change their own voice into the pretenders' voice. Speakers mimic the tone of voice that is relevant to the pretender, and concurrently express a particular attitude towards the proposition. Clark and Gerrig (1984: 122) describe this tone of voice as exaggerated, caricature and "heavily conspiratorial" (Clark and Gerrig, 1984: 122). As a result, the speaker's tone of voice mocking someone he/she pretends to be may influence the listener's understanding of irony like in Example 2.12:

Example 2.12

> Sherlock and Watson set out for a journey to find the villain Profession Moriarty. On their way they make a stop in the gypsy village and stay overnight. The Gypsies invite for dinner and Sherlock and Watson eat hedgehog goulash for the first time. Sherlock is pretending to be a well-educated person, praising his gypsy friend Madam Simza Heron for a wonderful hedgehog goulash:
>
> **Sherlock Holmes:** This is one of the best hedgehog goulashes I've ever tasted. I can't remember tasting it anywhere else.
> **Dr. Watson:** When was the last time you had hedgehog goulash?

(*Sherlock Holmes: A Game of Shadows*, 2011)

In Example 2.12 Sherlock is pretending to be someone else then he really is, e.g., an eloquent gentleman or an expert in hedgehog goulash appreciating its great taste. Sherlock's caricature tone of voice supports Watson's identification of irony.

More recently, Clark (1996: 364) has provided a refinement to Pretense Theory, describing irony as "joint pretense". In this refined version, speakers pretend to be imaginary people in an imaginary situation or fictional world jointly preforming a "two-layer act of communication" intended at listeners. In the first layer speakers are themselves performing an act in the present moment, while in the second layer speakers are pretending to be someone else in the imaginary situation. As a result, speakers and listeners are jointly

pretending to activate different layers of meaning, i.e., the literal and intended meaning. I have illustrated the conceptualisation of "joint pretense" drawing on Example 2.12 described above in which Sherlock and Madam Simza Heron are involved in unision in two layers of "joint action" Clark (1996: 364). At layer 1, Sherlock is prising Madam Simza Heron for the best hedgehog goulash. At layer 2, he appears to be an expert in hedgehog goulash who is praising a chief in a restaurant for an excellent meal. To join in the pretense, listeners must share the intended ironic meaning with speakers (Utsumi, 2000).

2.3.2 Objections to the Pretense Theory

Although in their Pretense Theory Clark and Gerrig (1984) proposed an alternative explanation to generate and identify irony through pretense (e.g., Gurillo and Ortega, 2013), there are several problem areas identified with Clark and Gerrig's (1984) study of irony which make their theorisation not fully capable of supporting the present analysis of multimodal irony in the film text.

The first criticism levelled at the pretense theory pertains to its limited ability to account for various instances of irony, especially in these cases where speakers are unable to pretend to be someone else to convey ironic meaning. For instance, Clark and Gerrig's (1984) theory appears to be incapable of supporting these utterances where a speaker expresses an absurd meaning or abstract idea since it is not possible to convey an absurd or conceptual content through pretense (Wilson, 2006; Clark, 2013). An illustration of an absurd concept is clearly visible in the case of Example 2.13:

Example 2.13

> Sherlock and Watson fight in the cathedral to save the life of an innocent girl. When they catch the black magician before he attempts kill the girl, it turns out he is the villain Lord Blackwood. Watson and Sherlock arrest Lord Blackwood before Inspector Lestrade arrive to the cathedral. At this point, Sherlock says to Lestrade: "Impeccable timing, Lestrade."

(*Sherlock Holmes*, 2009)

In Example 2.13, Sherlock is ironically commenting on Inspector Lestrade's timing, but irony is not construed through pretense. As a result, "pretense is not a necessary property of irony" (Utsumi, 2000: 1782) and in some cases, like in the example above, it is impossible to pretend abstract concepts like time.

According to Clark and Gerrig's (1984) theory, in order to understand irony, listeners have to recognise when speakers are pretending to be someone else meaning that the presence of the audience is essential to relay ironic meaning. However, there are also cases when irony arises, although the audience is not involved (Currie, 2005). This argument can be presented by Example 2.14 chosen from the present data set:

Example 2.14

> Watson visits Sherlock after a longer absence. When he opens the door to Holmes'
> flat, he finds himself in a lush jungle filled with a number of palm trees and wildlife
> plants inside. Watson comments Sherlock's study room with a deadpan voice:
>
> **Watson:** Your hedge needs trimming!
>
> (*Sherlock Holmes*, 2009)

In Example 2.14, Watson is mocking Sherlock for keeping his wild jungle inside of the flat which appears ridiculous for Watson. Although the audience (Sherlock) is not in this scene Watson's irony still seems to arise.

Wilson (2006) further contends that the Pretense Theory of irony requires elements of mimicry and simulation of someone else's tone of voice or attitude to recognise irony through pretense. There are situations, however, when irony arises without pretending someone else's voice. Ironists often produce ironic utterances using their own voice or their words as a form of articulation and do not need to pretend to be someone else to convey ironic meaning. Wilson (2006) illustrates this view with Example 2.15:

Example 2.15

> After a terrible game, Peter says to Mary:
> Peter: "I almost won"
> Mary (talking to a friend sitting next to her using her usual voice): "He almost won"

In Example 2.15, Mary responses to Peter's utterance "He almost won" using the flat and low-key intonation commonly designated as the "ironic tone of voice" to convey irony. But Mary is not imitating anyone's voice and is not pretending to be someone else. She is using her own voice and the ironic effect has still been achieved. As a result, there are a number of instances, in which a speaker is unable to convey irony through pretense but ironic meaning can still be achieved. Clark and Gerrig's (1984) theoretical account of irony does not allow the analysis of those instances of multimodal irony in which irony is not conveyed through pretense, or where the presence or the audience is not necessary to achieve the ironic effect, as illustrated by the several examples discussed in Section 2.2.

Another reason why the pretense theory is only partially capable of guiding the analysis of multimodal irony is its contribution of non-verbal elements to the generation and retrieval of ironic meaning that appears to be constrained. Clark and Gerrig (1984) focus exclusively on one non-verbal component, i.e., tone of voice or intonation. The "ironic tone of voice" is only one mode of communication that construes ironic meaning in the multimodal context, and pretense theorists did not elaborate on how other non-verbal elements like facial expression, gestures or body movements may reinforce the generation and interpretation of irony.

40

Sperber and Wilson (2012) believe that pretense could work more effectively, if it was combined with the element of echo. Pretense theory is therefore believed to be an alternative theory of irony to the echoic account proposed and developed by Sperber and Wilson (1981, 1992) within the theory of relevance (Clark, 2013). In the next section, the main assumptions of the echoic theory of irony will be examined, a theory whose inventors characterise as more straightforward and theoretically simpler than a pretense approach (Clark, 2013).

2.4 Echoic Theory of Irony

This section sets out to introduce the echoic theory of irony developed within relevance theory (Sperber and Wilson, 1992, 1995, 2012). Sperber and Wilson put forward their echoic theory of irony as an attempt to supersede Clark and Gerrig's pretense theory (1984) to analyse how irony is generated and interpreted. This section will evaluate the use of the central concepts developed within the echoic account of irony to inform this present study of multimodal irony.

2.4.1 Relevance theory and irony

The main aim of the theory of relevance put forward by Sperber and Wilson (1986, 1995) is to provide a comprehensive account of communicative processes in order to explain how communication can be (un)successful depending on how (in)direct speakers are.

It is necessary to clarify here what is meant exactly by the technical term "relevance". Sperber and Wilson (2012: 38) defined relevance as a "property of inputs to cognitive processes" which is determined by the amount of the processing efforts (i.e., the mental activities involved in comprehending a stimulus) and the cognitive effects (i.e., mental abilities in a cognitive system which modify the individual's representation of the world by drawing conclusions on the grounds of the existing or new assumptions) (Sperber and Wilson, 1995). The notion of relevance as a property of sounds, sights, memories or thoughts, among others, can therefore be described as a balance between cognitive effects and processing efforts (Sperber and Wilson, 1986, 1995; Clark, 2013). Sperber and Wilson (1992: 68) thus explain the term "principle of relevance" as "any utterance addressed to someone automatically conveys a presumption of its own relevance."

From a relevance-theoretic perspective, the most important condition for listeners is to find an interpretation that will be consistent with the principle of relevance. According to Sperber and Wilson (1995) listeners should presume that a stimulus is maximally relevant,

if achieved by as many cognitive effects and as little processing efforts as possible. A stimulus should be therefore at least optimally relevant to be worth the listener's attention to interpret it following the pathway of the least amount of processing effort. Thus, the generation and comprehension of irony is guided by the principle of relevance following the presumption of adequate cognitive effects and minimal necessary processing effort (Sperber and Wilson, 1986, 1992, 1995).

2.4.2 Irony as echoic use

Building on relevance theory, Wilson and Sperber (1995) developed the echoic theory of irony in an attempt to guide speakers and listeners in the process of generating and interpreting irony.

In the first instance, Sperber and Wilson (1981) conceptualised the notion of irony as a "self-referential use of words or other linguistic expressions" (Sperber and Wilson, 1981; 1992: 57). In other words, a mention is an identical reproduction of the original words or content (i.e., literal interpretation) and involves the reference to the words or content itself. Thus, ironic utterances are considered as literal interpretations of an attributed utterance or thought. Sperber and Wilson (1981) further distinguish two types of mentions i.e. explicit mention and implicit mention and the latter is used to convey ironic utterances. Nevertheless, Sperber and Wilson found the definition of irony as mention too limited as it could account only for these instances of ironic meaning in which irony is treated as literal interpretation totally excluding non-literal interpretations.

Thus, Sperber and Wilson (1986, 1992: 65) abandoned the notion of echoic mention in favour of the echoic interpretation defining "echoic utterances as echoic interpretations of an attributed thought or utterance, and verbal irony as a variety of echoic interpretation." That is to say, speakers are echoing an interpretation of an original utterance or thought and concurrently are dissociating from the utterance or thought echoed, conveying an attitude of disapproval behind it. The understanding and identification of verbal irony as echoic interpretation thus depends on the recognition of speakers' attitude (e.g., disapproving, derogatory). Nevertheless, the echoed material that speakers could dissociate themselves from seems to be limited to literal and non-literal interpretations only.

In 1995, Sperber and Wilson therefore re-analysed the echoic interpretation as the echoic use to extend the range of possible interpretations of irony and argued that an utterance is echoic, and thereby ironic, when it achieves "most of its relevance not by expressing the speaker's own views, not by reporting someone else's utterances or thoughts,

but by expressing the speaker's attitude to views she/he tacitly attributes to someone else" (Sperber and Wilson, 1995: 272). Hence, the primary aim of echoic use is not to provide information about the content of an attributed thought but to convey an attitude towards an utterance or thought that they attribute to someone else (Sperber and Wilson, 1995). The notion of echo and attitude are thus considered the key concepts in the echoic theory of irony.

2.4.3 The concept of echo

In their echoic account of irony, Sperber and Wilson (1981) proposed the notion of "echo" as a technical term to allow speakers to attribute a proposition to someone else and express a certain attitude, i.e., associative (positive) and dissociative (negative), and thus their reaction towards the entertained proposition. On this account, irony generates "a variety of echoic utterances used to express the speaker's attitude to the opinion echoed" and thereby allows speakers to dissociate from the proposition echoed, and listeners to identify and interpret ironic utterances (Sperber and Wilson, 1992: 59). In Sperber and Wilson's (1995) view, utterances need to be echoic in order to be considered as ironic.

Sperber and Wilson (1981, 1992, 1995) distinguished a wide variety of the echoed material (e.g., a situation, an utterance, cultural norm etc.) that speakers can dissociate from in order to produce ironic utterances. Primarily, a speaker echoes a proposition that is related to the immediately aforementioned utterance. In some cases, however, the proposition being echoed might have been produced some time ago or even in a very distant past (Sperber and Wilson, 1981). Speakers can also echo a representation attributed to a particular person (e.g., someone famous) or to a particular type of a person (e.g., a geek) to refer, for instance, to someone else's behaviour or traits of character in an ironic way. Moreover, ironic utterances can echo belief, hope, expectation or cultural norm (Sperber and Wilson, 1981, 1992; Sperber, 1984). This can be shown briefly in Example 2.16, in which the speaker (Watson) echoes a cultural norm in order to produce ironic utterance about a situation occurred:

Example 2.16

> Sherlock and Watson attend a masked ball with hope to find the villain Professor Moriarty. To avoid standing out of the crowd, Sherlock holds out his hand towards Watson and Watson comments when they start to dance together:
>
> **Doctor Watson:** I thought you'd never ask.
>
> (*Sherlock Holmes: A Game of Shadows,* 2011)

In Example 2.16, Watson intends to be ironic when Sherlock invites him to dance. He does so by echoing a culturally accepted norm (i.e., the formulaic form of acceptance), whereby a man conventionally asks a woman to dance, but not another man.

On this account, Sperber and Wilson's notion of echo will be employed as a conceptual tool to resonate with my analysis of multimodal irony. The application of the concept of echo allows the classification of the material from which an utterance is echoed as well as the isolation of moments in the scenes in which dissociative attitudes are expressed by speakers, thereby identifying the instances in which irony plays a pivotal narrative role in the selected scenes of the audiovisual text under study.

2.4.4 The concept of attitude

In their theory of irony, Sperber and Wilson (1995) gave primacy to the speaker's attitude vis-à-vis a thought, rather than the content of this thought. The notion of attitude is considered a form of expression to the proposition echoed that has less to do with the way the speaker is saying something and more how the speaker feels about what he/she is saying. The way the speaker feels about the proposition being expressed thus determines the generation and interpretation of irony.

From a relevance-theoretic perspective, verbal irony predominantly conveys an attitude of disapproval ranging from a mild disagreement, scepticism or mockery to the expression of a heavy derogatory or severe scornful and wry behaviour which Sperber and Wilson (1981, 1992, 1995) describe as "a tacitly dissociative attitude". An example of Sperber and Wilson's (1992) concept of attitude can be illustrated in Example 2.17, in which Mrs. Hudson expresses her dissociative attitude to the proposition echoed (i.e., Sherlock's utterance):

Example 2.17

Sherlock stays in the apartment for several weeks now designing and testing new weapons, costumes and poisons. Watson's maid Mrs. Hudson takes care of house and Sherlock, although she is not particularly fond of Sherlock. Once morning Mrs. Hudson brings Sherlock a cup of tea and says:

Mrs. Hudson: Tea, Mr. Holmes.
Holmes: Is it poisoned, Nanny?
Mrs. Hudson: There's enough of that in you already.

(*Sherlock Holmes*, 2009)

In Example 2.17, the echo here is the anaphoric reference to "poison" when Mrs. Hudson expresses an attitude of dissociation towards Sherlock in order to generate an ironic utterance. Drawing on Sperber and Wilson (1992), Sherlock's utterance "is ironical BECAUSE it is echoic: verbal irony consists in echoing a tacitly attributed [...] utterance

with a tacitly dissociative attitude" (Sperber and Wilson, 1995: 274; emphasis as in original). The manifestation of a certain sort of attitude of disapproval to a proposition or material echoed preponderantly involves an expression of a wide array of non-verbal semiotic resources. In particular, Sperber and Wilson (1981, 1992, 2012) draw special attention to the ironical tone of voice and different forms of kinetic behaviour that might facilitate the generation and interpretation of irony.

Audible components include the use of a peculiar tone of voice in the generation of irony. Not all ironical remarks employ a particular tone of voice, yet the ones that do help the audience identify the ironical intention (Ackerman, 1983; Rockwell, 2000; Bryant and Fox-Tree 2005). An ironic tone of voice is thus distinguished by "a flat or deadpan intonation, slower tempo, lower pitch and greater intensity that are found in the corresponding literal utterances" (Sperber and Wilson, 2012: 10). The generation and interpretation of irony may thus depend of audible components such as intonation, accent, timbre and volume of voice, to just name a few.

The attitude may also be expressed through a combination of other non-verbal elements involving visual or kinetic components. Sperber and Wilson (2012) argued that speakers may reinforce the intended ironic meaning and enhance irony comprehension using facial expressions, gestures or head movements, including "a wry facial expression, a resigned shrug, a weary shake of the head" (Sperber and Wilson, 2012: 1). Non-verbal elements are thus particularly helpful when listeners are unable to uncover the intended ironic meaning using background information, e.g., prior utterances, knowledge or contextual clues.

As Sperber and Wilson's notion of attitude involved a wide array of non-verbal components to express ironic meaning, it will be the central concept to guide this study of multimodal irony. The concept of attitude will thus allow the determination of the role that verbal and non-verbal elements play in the construal of irony in audiovisual texts and how they are intertwined to convey ironic meaning on screen.

2.4.5 Objections to the echoic theory of irony

Sperber and Wilson's echoic theory of irony has also been subject of criticism. Mostly critics have maintained that it is unable to account for all types of ironic utterances (Utsumi, 2000).

For instance, Seto (1998) argued against Sperber and Wilson's main claim that irony is necessarily echoic and argues that there are also cases of non-echoic irony. Seto (1998)

defines non-echoic irony as semantic reversal[12] associated with rhetorical devices such as oxymoron, litotes, paradox or hyperbole. For example, Seto (1998: 249) illustrated hyperbole as the instance of irony which is shown in Example 2.18 selected from the current analysis (my emphasis):

Example 2.18

> Watson and his wife Mary travel to Brighton to enjoy their honeymoon. Unexpectedly, troops of the villain Professor Moriarty attack them on the train. Sherlock walks into Watson and Mary's carriage to help them out. As the train goes very fast, Sherlock counts and suddenly throws Mary from the train to save her life, when Watson is busy shooting the officers. Mary lands in the river where Sherlock's brother Mycroft is waiting to rescue her. When Mycroft sees Mary is swimming in the river, he screams and comes closer to lend her a hand:
>
> **Mycroft Holmes:** Over here, Madam! I believe congratulations are in order. I'm the other Holmes.
> **Mary Watson:** You mean there's two of you? Oh, how *marvellous*! Can this evening be any better?

> (*Sherlock Holmes: A Game of Shadows,* 2011)

In Example 2.18 in response to Mycroft's introduction "I'm the other Holmes", Mary Sherlock used hyperbolic positive word "marvellous" to intensify her dissociative attitude towards both Holmes.

The echoic irony, on the other hand, according to Seto (1998, p. 249) is accompanied by linguistic markers such as certainly, definitely, evidently, indeed, truly, real(y) as illustrated by Example 2.19 chosen from the present data set:

Example 2.19

> Watson and Sherlock spend some time in the Gypsies' village before they start their journey to look for Professor Moriarty's factory. Sherlock is wearing now Gypsy clothes that he received as a gift and Watson mocks his new appearance saying:
>
> **Doctor Watson:** However, you do make a fantastic Gypsy.
> **Sherlock Holmes:** *Certainly,* smell like a fantastic Gypsy.

> (*Sherlock Holmes: A Game of Shadows,* 2011)

In Example 2.19 Watson is mocking Sherlock's outfit with the expression "a fantastic gypsy". Sherlock echoes Watson's expression with the linguistic marker "*certainly*" and thus reinforcing the ironic effect. Sperber and Wilson's echoic account of irony acknowledges the instances of irony when ironic utterances are strengthened by rhetoric devices and linguistic markers only to a limited extent.

---

[12] Sinclair (2004: 134) explains the notion of semantic reversal as follows: "Situations frequently arise in texts where the precise meaning of a word of phrase is determined more by the verbal environment than the parameters of lexical entry. Instead of expecting to understand a segment of text by accumulating the meaning of each successive meaningful unit, here is the reverse: where a number of units takes together create a meaning, and this meaning takes precedence over the 'dictionary meanings' of whatever words are chosen".

Similarly, Hamamoto (1998) contended that Sperber and Wilson's echoic theory is insufficiently explanatory as it fails to explain all possible instances of irony. For instance, their framework does not account for ironical utterances which describe existing situations or conditions (e.g., "Friends are always there where they need us") or cases when a literal meaning seems to have a negative appearance but aims to convey a positive meaning such as a wife can say "You are so naughty" to her husband after receiving present from him (Hamamoto, 1998: 266).

Despite the above-mentioned limitations, Sperber and Wilson's echoic theory of irony has been chosen to inform this study of multimodal irony since Sperber and Wilson acknowledge the importance of non-verbal semiotic resources in the generation and interpretation of irony, in contrast to Grice's (1967/79) and Clark and Gerrig's (1984) theory of irony. As a result, speakers may reinforce their intended ironic meaning by expressing their attitude to a proposition echoed through acoustic and visual components. On this account, along with the informative content of an utterance, speakers can support their reaction or attitude with a wide array of non-verbal elements.

2.5 Irony in Audiovisual Translation Studies

So far, this chapter has focused on the generation and interpretation of irony within the field of pragmatics. Having discussed three leading accounts of irony, Sperber and Wilson's echoic theory has been selected to explain the relay of ironic meaning in film translation. In this final section of this chapter Sperber and Wilson's attitude of dissociation has been incorporated into a proposition echoed in order to explore the importance of non-verbal components in the transfer of ironic meaning within Translation Studies, and more thoroughly within Audiovisual Translation. I also aim to examine the contribution of audiovisual translation scholars to the study of the translation of irony in audiovisual content.

In the 1990s, irony attracted relatively little scholarly attention within the domain of Translation Studies (Mateo, 1995). This may be explained by the untranslatability dogma frequently associated with the translation of irony (De Wilde, 2012). More recently, the study of irony has grown in importance and received significant critical attention from translation studies scholars (e.g., Chakhachiro, 2009; De Wilde, 2012; Feltrin-Morris, 2012). There are only a few scholars that have contributed to the study of irony within the field of Audiovisual Translation (Pelsmaekers & Van Besien, 2002; Zabalbeascoa, 2003). Thus, irony has remained a much-disputed subject of analysis, mainly due to cultural and linguistic discrepancies between the source and target languages.

The first problem that arises when attempting to analyse irony within Translation Studies is the very definition of the concept itself. Translation theories have been principally focused on the strategies or techniques analysing how to translate irony, rather than how irony came into being (Jakobson, 1959; Nida, 1964; Baker, 1992; Vinay and Darbelnet, 1995). Translation and audiovisual translation studies (AVTS) scholars (e.g., Pelsmaekers and Van Besien, 2002; Mateo, 1995; Hatim, 1997) often resort to definitions of irony that have been previously put forward by literary critics (e.g., Muecke, 1969; Booth, 1974), linguistic philosophers (Searle, 1976) or pragmatics scholars (e.g., Grice 1967/89; Sperber and Wilson, 1995). To date, however, there has been little agreement on a clear-cut definition of irony in TS and AVTS. The reason for this is presumably triggered by the fact that the notion of irony depends on the type of text being translated (e.g., written or audiovisual) and on the type of irony under analysis.

Great care needs to be taken when translating irony since it serves a number of vital communicative functions that can trigger different effects in listeners/viewers (Pexman and Zvaigzne, 2004). For instance, irony allows to criticise through sarcasm, which involves saying something positive to convey a negative meaning, e.g., when saying "You look gorgeous today" to someone in disheveled clothes (Pexman and Zvaigzne, 2004). Sarcasm represents "an especially negative form of irony" (Gibbs, 1994: 384) or "the crudest form of irony" (Muecke, 1982: 54) evoking irritation to put someone down rather than make laugh (Chakhachiro, 2009). Irony also allows the speaker to mock or tease someone (e.g., Pexman and Zvaigzne, 2004) and to be humorous (e.g., Robert and Kreuz, 1995) with the help of linguistic drives such as "poetic references, paradoxes and puns" (Chakhachiro, 2009: 33) or parody. Here, parody can be described as "an ironic interpretation of an intertextual relation – between two texts […] between a text and a genre or between a text and a style or a convention" in which a speaker is entertained at the cost of a specific person, work of art, institution or convention (Elleström, 2002: 153). It is therefore highly important for a translator to identify the function irony performs in a text/film.

Studies of irony also show other problem areas related to the translation of irony, i.e., transfer of cultural references in ironic utterances. Some scholars (Mateo, 1995; Pelsmaekers and Van Besien, 2002) contend that it is difficult to find appropriate equivalent phrases in the target culture that trigger similar reactions to the ones that ironic utterances in the source language (SL) evoke among their readers, as irony can be deeply rooted in the source culture (Feltrin-Morris, 2012; Hatim and Mason, 1990) "we cannot expect that all cultures have similar understandings and uses of irony" (Barbe, 1995: 144). Similarly, sometimes the

context may be crucial to understand irony, if SL writers and TL readers share the contextual clues (Yus Ramos, 1997/98). Even though, the translation of irony is a challenge, it should not be omitted in the TL, as the intended meaning could otherwise be misinterpreted (Barbe, 1995). Therefore, translators have to prioritise the transfer of irony in texts/films in which ironic utterances are essential to interpret the translation correctly (De Wilde, 2012).

In contrast to the translation of irony in print products like books or newspapers in which the verbal message is primarily transferred through writing, the translation of verbal irony in audiovisual products is transferred through four modes of communication, i.e., "dialogue, music and effects, picture, and — for a smaller part — writing (displays and captions)" (Gottlieb, 2004: 86). Audiovisual translation scholars contend that verbal and non-verbal elements are intertwined in different ways to produce "different patterns of cohesion, intertextuality and the other features of textual structure and meaning" (Zabalbeascoa, 2008: 23). Irony in audiovisual texts seems therefore to be more translatable than in written texts as its transfer is not limited to verbal or contextual components only but it can be supported by the combination of verbal and non-verbal constituents.

Zabalbeascoa (2003) explores the notion of "audiovisual irony" and its translation in *Transpotting* (1996) drawing on the premise that verbal and non-verbal sign systems, audio and visual elements are intertwined to convey meaning in audiovisual texts. Within the category of non-verbal irony, Zabalbeascoa (2003) paid particular attention to visual (e.g., the picture, photography) and audio (e.g., music) components. He analysed several scenes in which irony is conveyed non-verbally rather than verbally. By way of illustration, in Example 2.20 Zabalbeascoa (2003: 305) presented audiovisual irony through Renton - one of the characters of the film *Transpotting*, a film in which visual components play a pivotal role in the construal of ironic meaning:

Example 2.20

> **Renton** (voice-over): "Choose life. Choose a job. Choose a career. Choose a family. Choose a fucking big television. Choose washing machines, cars, compact disc players, and electrical tin openers." Suddenly, as Renton crosses a road, a car skids to a halt, slightly knocking him. In a moment of detachment, he stops and smiles at the shocked driver.

In Example 2.20, Zabalbeascoa (2003) considered Renton's smile ironic because the character does not have any reasons to be happy in the context of the situation described above. Zabalbeascoa (2003) argued further that there is a wide range of audio components that may enhance ironic meaning, for instance, laughter, screams, knocks, voice imitations as well as music and special effects.

Zabalbeascoa (2003) indeed explored different patterns of verbal and non-verbal sign systems aimed to convey ironic meaning. However, a more detailed study involving multimodal transcription, which enables a detailed dissection of film texts, will allow me to explore more thoroughly how the on-screen characters express their dissociative attitude to a proposition echoed through the simultaneous composite of a wide array of non-verbal semiotic resources. Drawing on Baldry and Thibault's (2006) concept of "resource integration principle" and their understanding of multimodality, it will be possible to explore which non-verbal semiotic resources participate concurrently in the construal and relay of irony in audiovisual texts.

Pelsmaekers and Van Besien (2002), on the other hand, interrogated the phenomenon of humorous irony in film subtitles. Their study set out to examine how subtitlers managed to retain a certain degree of ironic meaning relayed into subtitles. For the purpose of their study, Pelsmaekers and Van Besien (2002) conducted a qualitative analysis of 211 humorous ironic utterances with Dutch subtitles from 12 episodes of *Blackadder.*

In contrast to Zabalbeascoa's (2003) taxonomy, Pelsmaekers and Van Besien (2002) chose a pragmatic perspective to inform their study of irony. Drawing on Searle's (1969/76) theory of speech acts, Pelsmaekers and Van Besien (2002: 243) attempted to elaborate their definition of irony, according to which says that irony arises when "there is some kind of contrast or incongruity between what is said (the propositional content) and what can be inferred from the situation".

The results of their study indicate that ironic potential was retained in a vast majority of Dutch subtitles, although in two-third of the cases verbal ironic cues were modified and deleted. Nevertheless, the modification and deletion of certain verbal cues such as address forms, intensifiers, hedges and echoes of previous utterances uncovered another form of ironic utterances conveying deadpan irony, i.e., when speakers appear to be serious and they are hiding the fact that they are ironically joking at someone (Haiman, 1990).

In their analysis and translation of irony, Pelsmaekers and Van Besien (2002) also highlighted the interplay between verbal and non-verbal components. They observed that listeners should combine verbal and non-verbal ironic cues in order to indicate that speaker's dissociative attitude is not taken literally or positively. According to Pelsmaekers and Van Besien (2002: 245) speakers may use non-verbal signs like "nasalisation, slow and emphatic speech, snorting sounds and marked intonation". On this account, the integration between subtitles and non-verbal signs on screen is essential. Pelsmaekers and Van Besien (2002)

further argue that subtitles cannot be read in isolation from sounds and pictures as otherwise the humorous ironic effect might be lost.

Nevertheless, Pelsmaekers and Van Besien's (2002) analysis of non-verbal elements in the creation of subtitles is still very limited. Admittedly, they note the importance of non-verbal components in the translation and interpretation process, yet a far more detailed study focusing on the contribution of non-verbal constituents to irony construal is necessary in order to explore how verbal and non-verbal modes participate in the transfer of ironic meaning in audiovisual productions. Furthermore, although Pelsmaekers and Van Besien (2002) noticed that the paralinguistic ironic cues may enhance the understanding of irony, they do not provide any empirical evidence of how non-verbal semiotic resources contribute to facilitate the comprehension of irony by viewers. Here, an experimental study involving eye tracking technology and the administration of questionnaires will allow the exploration of the verbal and non-verbal semiotic resources and their contribution to irony comprehension in both subtitled and voiced-over films.

2.6 Conclusions

This chapter has focused on the study of irony within the domain of pragmatics. A critical overview of three main theoretical approaches in the theorisation of irony has been provided, i.e., Grice's theory of conversation (1967/89), Clark and Gerrig's pretense theory (1984) and Sperber and Wilson's echoic theory (1981, 1992, 1995) in order to see which theoretical framework allows the analysis of multimodal irony in the most comprehensive manner. Subsequently, the most recent studies of irony in the field of audiovisual translation (Pelsmaekers and Van Besien, 2002; Zabalbeascoa, 2003) have been discussed, accentuating the importance on non-verbal components in the transfer of ironic meaning on screen and identifying the gaps that should be addressed.

The analysis of three leading pragmatic-driven theories of irony in this chapter has revealed that the theoretical apparatus of irony developed by Sperber and Wilson's (1981, 1992, 1995) echoic theory of irony should provide a sufficiently strong basis to support the analysis of multimodal irony, while Grice's and Clark and Gerrig's account of irony are not fully capable of informing the present study.

Although Grice's contribution to the explanatory theory of irony in pragmatics is considered as very important (Clark, 2013), a number of shortcomings have been identified in his approach with regard to the present study. First and foremost, Grice was primarily focused on irony conveyed verbally and paid far too little attention to the non-verbal

dimension of irony. He admittedly observed that a particular attitude or tone of voice may help to generate and interpret irony. However, Grice did not explain these two aspects of irony any further and did not provide any examples to support his contention. Due to the limited analysis of non-verbal constituents in the construal of irony therefore, Grice's theory of irony is not capable of informing the study of multimodal irony, in which both verbal and non-verbal components play a pivotal role.

On the other hand, Clark and Gerrig's pretense theory (1984) is only partially capable of guiding the analysis of multimodal irony. Although Clark and Gerrig (1984) took into account non-verbal components in the construal of their theory of verbal irony, they essentially focused on audible elements, such as tone of voice or intonation, which is only one of the components contributing to relay of ironic meaning. Moreover, as noted in section 2.2, it has been argued by several scholars (e.g., Utsumi, 2000; Wilson, 2006) that Clark and Gerrig's (1984) notion of pretense is not enough to generate and interpret different instances of irony and its functions. Sperber and Wilson (2012) believe that pretense could work more effectively if it was combined with the element of echo within their own conceptualisation of irony.

In comparison with the account of irony proposed by Grice (1967/79) and Clark and Gerrig (1984), the echoic theory of irony developed by Sperber and Wilson (1981, 1992, 1995) appears to be the most adequate framework to support the analysis of multimodal irony in audiovisual texts due to two main reasons. First, the concept of echo accounts for a wide variety of ironic utterances and allows the description of how listeners identify and interpret irony. Second, the concept of attitude may reinforce the intended ironic meaning through the combination of verbal and non-verbal semiotic resources. As the preceding discussion in subsection 2.4.4 highlights, in their echoic view of irony, Sperber and Wilson bring to the foreground the significance of non-verbal components in shaping of the notion of irony and irony comprehension. Nevertheless, several limitations of Sperber and Wilson's account have also been identified in relation to the present study of multimodal irony.

First, it has been argued that Sperber and Wilson's account of irony is not fully adequate to explore the construal of irony in audiovisual texts. Building on relevance theory, the echoic theory primarily aims to analyse the generation and interpretation of irony in everyday communicative interactions between speakers and listeners rather than between the on-screen characters and the viewers of audiovisual productions. The combination of Sperber and Wilson's operationalisation of irony with multimodal theory would thus

facilitate the analysis of irony construal and comprehension in multimodal texts like films on a vertical level, i.e., between the on-screen characters (Vanoye, 1985).

Second, earlier studies of irony in the field of pragmatics (i.e., Grice, 1967/79; Clark and Gerrig, 1894) predominantly focused on the mechanism of verbal irony, similarly to Sperber and Wilson's (1981, 1992, 1995) theoretical framework. Nevertheless, Sperber and Wilson acknowledge the importance of non-verbal elements in the generation and interpretation of verbal irony. Their concept of attitude conveys ironic meaning in a multimodal rather than monomodal way as "the synergistic construction of meanings […] cannot be derived from either mode separately" (Unsworth and Cleirigh, 2009: 150). The speaker is therefore expected to convey ironic meaning through the attitude concurrently combining different communicative modes.

However, Sperber and Wilson fail to explain what these non-verbal semiotic representations are that speakers convey through their expression of attitude and how these semiotic resources are simultaneously intertwined to create a meaningful whole. In what follows, it is crucial to explore to what extent non-verbal semiotic resources contribute to the construal and comprehension of ironic meaning. Given Sperber and Wilson's conceptualisation of irony, it appears essential to incorporate a new framework developed within the multimodal theory to examine the contribution of non-verbal semiotic resources to irony construal and comprehension in audiovisual productions.

The study of irony on screen from a multimodal approach would also represent an important contribution to Audiovisual Translation Studies. Although, more recently, research on verbal irony has been attracting growing attention from audiovisual translation scholars (e.g., Pelsmaekers and Van Besien 2002, Zabalbeascoa 2003), the number of studies on the amalgamation between verbal and non-verbal components in the analysis and translation of irony in the subtitled version of audiovisual texts is still very limited, in particular with regard to the voiced-over translations of films, on which to the best of my knowledge, no study has been conducted so far. It is therefore important to conduct a more detailed analysis as to which and how non-verbal semiotic resources are combined multimodally to construe and transfer meaningful ironic meaning in the translation of audiovisual texts.

Although extensive research has been carried out on the study of irony in the field of pragmatics, few studies exist which can adequately determine the contribution of non-verbal constituents to the construal, translation and comprehension of irony in multimodal texts like films. Despite the above-mentioned weaknesses, the echoic account of irony proposed by

Sperber and Wilson has been chosen to inform the present study. Nevertheless, in order to be fully capable of analysing the notion of multimodal irony in audiovisual productions, Sperber and Wilson's conceptualisation of irony has to be combined with a multimodal account. The multimodal approach to the analysis of irony in films and in filmic dialogue will be thus thoroughly explored in the next chapter.

3 MULTIMODAL IRONY

3.1 Introduction

Chapter 2 examined the research on irony informed by pragmatics, which predominantly focused on identifying and describing its verbal features (Grice, 1967/89; Clark and Gerrig, 1984). The interplay between verbal and non-verbal cues in the generation and comprehension of ironic utterances has subsequently been explored by Sperber and Wilson (1981, 1992, 1995) in their echoic theory of irony. The contribution of non-verbal semiotics to the mediation of audiovisual texts has also been analysed by translation scholars (e.g., Pelsmaekers and Van Besien, 2002; Zabalbeascoa, 2003). What remains to be explored is the synergy of different non-verbal meaning-making resources in the construal of multimodal ironic content on screen. This chapter therefore incorporates Sperber and Wilson's echoic theory of irony as my theoretical framework for the analysis of how non-verbal elements in films typically express the speaker's dissociative attitude in the film text.

This chapter is structured as follows. Section 3.2 examines these non-verbal markers that support the generation and retrieval of ironic utterances. Section 3.3 explores irony as a modality of cultural expression in the films under analysis. Section 3.4 investigates how meaning, particularly ironic meaning, is multimodally construed with the help of non-verbal film language, such as mise-en-scène, cinematography, editing and sound. Section 3.5 concentrates on the narrative role of irony in film dialogue, while Section 3.6 discusses the notion of intertextuality of irony in the genre of action comedy. Finally, Section 3.7 proposes the use of the notion of "multimodal irony" to describe how irony is embedded across multiple semiotic resources in an audiovisual environment.

3.2 Non-verbal Markers of Irony

Although several theorists have already examined the verbal features of irony (e.g., Grice, 1967/89; Clark and Gerrig, 1984), more recently, a growing number of researchers have shifted their attention to the non-verbal properties of irony encapsulated in the speaker's dissociative attitude (e.g., Rockwell, 2000; Attardo et al., 2003; Bryant and fox Tree, 2005; Gibbs, 2011).

Several scholars (e.g., Sperber and Wilson, 2012; Kreuz and Roberts, 1995) have noticed that speakers tend to reinforce their ironic attitude of dissociation using acoustic markers in their communicative interactions, in particular when the contextual knowledge shared by the speaker and the listener is limited. The acoustic features of irony are primarily

represented by "the ironic tone of voice" which involves the use of different voice modulations (Bryant and Fox Tree, 2002: 110). The results yielded from empirical studies on the ironic tone of voice reveal that "naturally produced prosody can significantly contribute to listeners' recognition of verbal irony" in spontaneous (Bryant and Fox Tree 2002: 110) and in prefabricated speech (e.g., Anolli et al., 2000; Rockwell, 2000). The contribution of acoustic cues to irony comprehension has also been confirmed in languages other than English, i.e., German (Scharrer and Christmann, 2011) or Italian (Anolli et al., 2000). Among the various properties of voice quality, intonation appears to be the most prevailing index of ironic content (Attardo et al., 2003). It is thus realised through a specific intonational pattern that can be characterised as "relatively monotonous" (Haiman, 1998: 39). In the majority of utterances produced in English, ironic intonation is identified as flat (i.e., neither rising nor falling, Haiman, 1998); and only in the case of questions, as rising (Schaffer, 1982). Other characteristics of pitch modulations are examined, for instance, by Antolli et al. (2000) who revealed that a lower pitch distinguishes ironic voice. In contrast, a few years later, Rockwell (2000) recognised a higher pitch as an indicator of irony. Several researchers have also speculated whether irony can be detected through "slowed speech rate, prolonged syllables" (Bryant and Fox Tree, 2002: 101; Haiman, 1998) or heavy stress (Gibbs and O'Brian, 1991). In other cases, pauses or extra-long pauses between words have been considered as markers of irony (Haiman, 1998; Schaffer, 1982). Finally, laughter syllables[13] (Haiman, 1998) and laughter (Jefferson, 1984) in general have also been identified as markers of ironic content underscoring the humorous features of irony. Bryant (2011: 294) also noticed that "antiphonal laughter"[14] can signal the presence and understanding of irony.

Apart from the analysis of acoustic cues, relatively fewer "non-verbal researchers" (Rockwell, 2000: 484) have examined the contribution of the speaker's kinesic body language to the generation and retrieval of irony. The kinesic actions thus deserve particular attention as they appear to be more informative than intonational cues in the interpretation of irony (Kreuz and Roberts, 1995). For instance, Attardo et al. (2003) revealed that several features of facial expression such as eyebrows (raised, lowered), eyes (wide open, squinting, rolling), winking (cf. Muecke, 1978) or smiling help to convey and grasp ironic meaning. Attardo et al. (2003: 254) also put forward the notion of "a blank face" which is defined as "expressionless, emotionless and motionless". As a result, "a blank face" avoids any facial

---

[13] Laughter syllables are repetitive laughs that may sound like "ha-ha-ha" or "ho-ho-ho" and "have a strong harmonic structure, a multiple of a low (…) frequency" (Andrews, 2013: 147).

[14] Antiphonal laughter can be defined as "laughter between social partners that occurs in close temporal proximity" (Wheeler, 2013: 88)
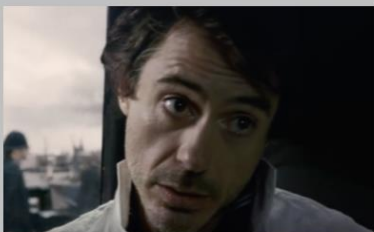
movements; in other words, it involves "no smile, no grimace, no eyebrow raising, no frown". In further studies, Williams (2009) suggests that the direction of gaze may also indicate ironic intent, especially when speakers drift their gaze away from an interlocutor.

As an illustration of acoustic and kinesic markers of irony, consider the following scene in Example 3.1 from *Sherlock Holmes* (2009) in which Sherlock expresses his attitude of disapproval towards Dr. Watson.

Example 3.1

Watson and Sherlock are travelling in a carriage towards the meeting with Inspector Lestrade. Watson still bears a grudge towards Sherlock for his behaviour towards his future wife Mary in a restaurant. Despite Sherlock's attempts to start a conversation, Watson still does not say anything. Finally, Sherlock comments on Watson's silence:

Frame No	Time (h:m:s)	Visual frame	Film dialogue	Non-verbal markers of irony
1	00:22:09		*You have the grant gift of silence, Watson*	Sherlock pulls back his head on the right side looking on Dr Watson from above. Sherlock's face is emotionless typical for 'blank face'. Sherlock's dialogue is expressed with flat intonation using a lower pitch.
2	00:22:10		*It makes you quite invaluable*	Sherlock raises his eyebrows and keeps eyes wide open. He wrinkles and keeps his head straight. Sherlock's dialogue is expressed with flat intonation using a lower pitch.
3	00:22:10		*as a companion.*	Sherlock turns his head on left, raises his eyebrows, wrinkles and keeps his eyes wide open with a characteristic smirk. Sherlock's dialogue is expressed with flat intonation using a lower pitch.

In Example 3.1, Sherlock expresses his attitude of disapproval with a wide range of non-verbal markers, thus reinforcing his attitude of disapproval towards Watson. Sherlock's emotionless "blank face", his raised eyebrows, wide opened eyes and looking at Dr. Watson

from below are vital in order to mock his companion. Sherlock's body language in unison with a flat intonation and a lower pitch illustrate typical expressions of ironic intent according to Haiman (1998) and Antolli et al. (2000). Yus (2000) argues that non-verbal markers like those mentioned in Example 3.1 can form an ancillary part of the generation and interpretation of ironic utterances. That is to say, the contextual cues such as "the speaker's ostensive tone of voice and facial expression, [as well as] the visual information" are particularly helpful when the listeners are unable to uncover the intended ironic meaning (Yus, 2000: 3). Yus further contends that it is easier to retrieve ironic meaning when non-verbal resources accompany verbal irony and identifies "non-verbal behaviour as an explicit irony marker" (2000: 3).

While the above discussed studies examine the comprehension of non-verbal markers of irony by people from different linguistic and cultural backgrounds no study, to the best of my knowledge, has been carried out so far to understand the impact of non-verbal behaviour on irony comprehension by Polish viewers, or the extent to which these non-verbal markers are determined by culturally-oriented conventions.

Furthermore, the studies discussed in this section largely explore non-verbal markers of irony in spontaneous (Gibbs, 2011; Bryant and Fox Tree, 2002) and pre-fabricated communicative interactions (Scharrer and Christmann, 2011; Rockwell, 2000; Ackerman, 1983). Very few studies, however, have so far investigated the non-verbal nature of irony in films, particularly in those that expose audiences to a phenomenon like irony – that, in this case, is presented in the context of recent adaptations of well-known stories about the British detective, Sherlock Holmes, as a new representation of his character.

In the following section, a Cultural Studies approach is explored as a tool to understand how irony is marked and used in Guy Ritchie's Sherlock Holmes films (2009, 2011) to discuss the limitations mentioned above.

3.3 Multimodal Irony as a Mode of Cultural Expression in Sherlock Holmes Films

Even outside of his British homeland, the traditional characteristics of the British cultural icon Sherlock Holmes such as a pipe, magnifying glass or deerstalker, as described in Sir Arthur Conan Doyle's stories have become well known. These signs have become well-rooted trademarks of Sherlock Holmes' cultural identity within the British society since the late nineteenth century. The more recent adaptations, however, present the consulting detective as a more universal and modern character embodying new trademarks such as an

ironic attitude of dissociation that foreign language viewers may find difficult to interpret since "irony is deceptive and can't be conveyed easily across cultures" (Pugliese, 2010: 45). Before analysing Ritchie's interpretation of Sherlock Holmes' films in which irony plays a pivotal role, this section will examine the notion of culture and cultural representation as they have been discussed within the culture studies domain.

According to Hall (1997: 1), the concept of culture has traditionally been conceptualised as being "the best that has been thought and said in a society" as represented by "high culture" and expressed in the form of classical paintings, literature or philosophy. In more "modern" approaches, the concept of culture encompasses a wider range of forms of art, design, music or mass-produced entertainment known as "popular culture" that defines "a way of life" of large audiences worldwide (Hall, 1997).

Culture does not purely consist of a set of paintings, comics, or films. Instead, it is primarily "concerned with the production and the exchange of meanings - the giving and taking of meaning - between the members of a society or group" (Hall, 1997: 2). That is to say, people belonging to the same culture are expected to interpret and express their thoughts, emotions as well as irony in roughly the same way, so that they can understand each other. There is, however, a certain degree of diversity in the way meaning is represented and interpreted in any cultural circle. It is also crucial to understand how language uses the functions of non-verbal forms of expression such as sounds, images or objects to construe, represent and transmit meaning, in particular ironic meaning.

Representation is thus a key concept here. It is fundamental to the generation and exchange of meaning process across all members of a culture (Hall, 1997). To understand how language, sounds and images are used to represent the world and irony meaningfully, Hall (1997) elaborates on three different approaches to representation, i.e., the reflective, the intentional and the constructionist.

In the reflective approach, a representation of objects, ideas, events or language functions reflects a true meaning that exists in the world. In order to have a successful conversation, people need to share the same concept of a word in their own cultures, otherwise the communication will be misunderstood (Hall, 1997). Since the notion of irony may vary from culture to culture, it may be challenging for participants to share the same concept of irony.

The intentional approach holds that the speaker uses language to impose his/her intended meaning on the world (Hall, 1997). For instance, question tags may be interpreted as displays of irony by certain viewers.

The constructionist approach argues that meanings are construed within and across cultures with the use of signs (Hall, 1997). Meanings are therefore construed rather than pre-existent within cultures. Constructionist theorists distinguish between the material world and the symbolic world. The material world comprises of people and actions or events, while the latter includes representational tools such as meaning and language. The material world cannot produce meaning without using representational systems like signs from the symbolic world (Hall, 1997). The sign is thus essential to construe meaning, including ironic meaning, in a culture as all cultural objects and practices make use of signs to convey meaning (Hall, 1997). In a culture, however, "meaning often depends on larger units of analysis - narratives, statements, groups of images, whole discourses which operate across a variety of texts […]" (Hall, 1997: 42). Thereby, words are often not sufficient and other forms of representations have to be taken into account. For instance, the image of Sherlock Holmes comprises of different signs which are then transferred across other cultures without using verbal language.

The co-deployment of diverse signs that participants in a culture use to create representations for objects, people or events produce meanings that give people a sense of their own identity and a feeling of belonging to a certain group (Hall, 1997). However, there is no "one, single, fixed and unchanging meaning" as meaning is constantly evolving (Hall, 1997: 4). According to Hall (1997) in everyday life, as part of social interactions within and between cultures people continuously produce and exchange meanings, including ironic meaning, as a result of the modern means of communication through the mass media technology. For instance, the characteristic figure of Sherlock Holmes has also gained new representations since it was first created by Conan Doyle in the late nineteenth century, as a result of technological advances that have made the stories of the consulting detective accessible worldwide.

In the 21st century, Sherlock Holmes' stories have evolved and inspired producers of many different genres such as films, TV-series, comics or radio programmes who have come up with new representations of the detective. In the context of this thesis, I focused on the two film adaptations under analysis, i.e., *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011) in which the director Guy Richie and the actor Robert Downey Jr. transformed and modernised Sir Arthur Conan Doyle's character. Although this representation contains symbols and signs from the original source material so that fans are not disappointed, they also reveal novel characteristics of Sherlock Holmes, "balancing otherness and familiarity" (Di Giovani, 2003: 215).

In contrast to Conan Doyle's portray of Sherlock Holmes, Guy Richie attempts to reinvent the character of Holmes and to present a more universal and diverse image of the detective (Ritchie, 2009). Ritchie abandoned the clichés surrounding Sherlock and incorporates totally novel trademarks appealing to large audiences worldwide in the 21st century (Jensen, 2014). Given the considerable cultural changes that have happened since the traditional trademarks of Sherlock Holmes were first seen, Ritchie attempted to extend the boundaries of a single nation and to demonstrate the diversity and universality of Sherlock Holmes.

To convert the British character into a universal icon, several strategies have been employed on both the linguistic and cultural level. Although in Ritchie's films Sherlock inhabits Victorian London, cultural references have been adapted and a number of elements and expressions have been modified to fit the contemporary culture within the modern world (Di Giovani, 2003). On this account, in order to reflect the ideas and values of the current world, referring to widely spread elements "becomes a real necessity when aiming to appeal large groups and mass audiences" (Di Giovani, 2003: 212).

In her analysis Jensen (2014) distinguishes novel and more contemporary trademarks that reveal the universal character and values of Conan Doyle's figure and describes the detective as an action hero and a comedian. Jensen (2014) notices that "an action-hero carries meanings for people in many different parts of the world" which is a possible reason for Ritchie to fit Sherlock Holmes within these conventions. In Ritchie's representation of Holmes, Sherlock fights using martial arts against his opponents in order to rescue his dearest companion Dr. Watson or to rescue civilisation from the danger caused by Professor Moriarty and his nuclear weapon. To some extent, Sherlock can also be categorised as a super hero who is using the power of deduction and observation skills to defeat his opponent. Given the terrorist attacks and wars that have filled media headlines in the 21st century, the notion of the action and super hero has become a particularly significant and meaningful sign detected in films in many cultures (Shadrina, 2014).

Despite the fast and action-packed plot, the two Guy Ritchie-directed Sherlock Holmes films also incorporate comical and humorous aspects of Sherlock Holmes, predominantly in order to catch attention of and entertain as many viewers as possible (Di Giovani, 2003: 213). Nevertheless, for instance in Sherlock Holmes (2009), humour is largely achieved through Sherlock's ironic and dissociative attitude as he purposely displays his high intellectual superiority in the presence of others through his wordplay, banter and behaviour (Lyall, 2009; Scott, 2009). For example, on several occasions Sherlock expresses

this attitude of disapproval mocking the competence of Inspector Lestrade and revealing him to be ignorant. Sherlock was not presented as an ironic character in Conan Doyle's stories, and thus irony can also be considered new trademark that Ritchie added in his modernised representation of the detective.

Since irony is complex and often described a very culture-specific phenomenon (Barber, 1995), some audiences, including Polish, may experience some difficulties in interpreting Sherlock's attitude as ironic. Nevertheless, there are several elements that can facilitate the irony comprehension process among new audiences.

As described in Section 3.2, there are numerous non-verbal markers of irony, such as the speaker's vocal modulations or facial expressions, which can help viewers decode Sherlock's attitude of dissociation. Although the interpretation of the non-verbal elements conveyed visually or acoustically can differ in various socio-cultural environments (Antonijevic, 2008; Williams, Burns, Harmon, 2009), the set of non-verbal elements that often characterise irony in Ritchie's representation of Sherlock Holmes may help Poles understand that he is being ironic (see Section 6.4). These non-verbal elements of Sherlock's behaviour are also associated with certain types of emotions. People across cultures are equipped with a set of universal emotions that may be understood in a similar way. At the same time culture provides us with another range of unique and culture specific emotions that target viewers may acquire by means of "affective socialisation and immersion in a L2 culture and language" (Ożańska-Ponikwia, 2013). In this way, viewers may also extend their fundamental emotional repertoire and recognise some common feelings for the source and target culture.

Robert Downey Jr.'s representation of the detective also helps audiences interpret Sherlock Holmes' dissociative attitude as ironic through his non-verbal communication particularly when he intends to relay emotions with his eyes as "irony is a gift Downey likes to share with audiences" Burr (2011). The focus has so far remained on the individual non-verbal markers of irony, rather than on the way these markers are co-deployed to construe ironic meaning simultaneously in multimodal texts like films. In their thorough study, Attardo et al. (2003) reveal a range of non-verbal markers of irony that American viewers identified while watching selected clips from television situation comedies. Nevertheless, it is little known whether and/or to what extent the interaction of different elements contributes to the construal and interpretation of irony on screen. Therefore, in the following section, an inter-disciplinary approach has been applied to the analysis of ironic meaning in films to

examine how ironic intent is construed through various meaning-making resources, both verbal and non-verbal, in films.

3.4. Multimodal Approach to the Study of Irony

The multimodal research gradually shifted its focus from "linguistic analysis over the analysis of other modes" (Scollon and Scollon, 2014: 180) defining the terms "**mode**" and "**modality**" as different types of acoustic or visual meaning-making resources that are combined to create a meaningful whole (e.g., Chandler, 2002). Multimodality thus represents a relatively novel alternative approach to the study of communication and representation. In this view, language is no longer seen as an independent mode but rather as one of several modes embedded in multimodal communicative acts (e.g., Scollon and Scollon, 2014).

In the late 1990s, pioneering scholars Kress and van Leeuwen (1996/2006) started analysing visual and acoustic modes that contribute to the creation of meaning. In their social semiotic approach to multimodality, Kress and van Leeuwen (1996/2006) examined how semiotic resources are selected and designed for meaning-making by the individual's interest and motivation in certain social contexts. Research within their framework is concerned with resources such as gesture and image, exploring how they are configured and realised in different communicative situations. Kress and van Leeuwen (1996/2006) perceive semiotic resources as dynamic, context-dependent elements which create meanings in synergy with other resources.

Drawing on Kress and van Leeuwen's work (1996/2006), Baldry and Thibault (2006) broadened the conceptualisation of multimodality, defining it as the "diverse ways in which a number of distinct semiotic resource systems are both *co-deployed* and *co-contextualised* in the making of a text-specific meaning" (Baldry and Thibault, 2006: 21; empahsis as in original). This "diversity of meaning-making activities" is reflected by the combination of verbal and non-verbal modes that jointly contribute to the perception of multimodal texts as a whole. As a result, meaning does not arise from the integration of individual semiotic resources, but is construed through their co-deployment. The combination of different semiotic resources involving images, sounds and words is encapsulated in the notion of "resource integration principle" which is at the heart of Baldry and Thibault's approach to multimodality. The principle determines the ways modes jointly operate to create a whole. Baldry and Thibault (2006: 18) outline their standpoint in the following terms:

resources are not simply juxtaposed as separate modes of meaning making but are combined and integrated to form a complex whole which cannot be reduced to or explained in terms of the mere sum of its separate parts.

In Baldry and Thibault's (2006: 18) account, several resources combined together produce **multimodal texts** defined as "composite products of the combined effects of all the resources used to create and interpret them". In their study, Baldry and Thibault (2006) analyse different multimodal configurations in diverse multimodal texts like the printed page or web pages. This present thesis concentrates on the analysis of multimodal irony in a dynamic type of multimodal text: film.

3.4.1 Multimodal irony in the film text

Since the film text is considered "a complex symbol system, made up of images, words and sounds" interwoven on screen (French, 1981: 240), it is essential to come up with a structured and robust approach to the study of multimodal irony on screen.

To be able to account for a wide array of elements that contribute to the construal of multimodal irony in the film text, Baldry and Thibault's (2006) classification of modes of the film text, (e.g., visual image, kinesic actions and soundtrack) has been combined with the insights derived from the film studies literature. Hence, in order to keep consistency in terminology across chapter three these modes will be referred to as the visual, kinesic and acoustic modes, respectively. They are expressed by a combination of elements pertaining to the language of film, i.e., mise-en-scène, cinematography, editing and sound.

There is a need to clarify exactly what is meant by the "visual", "kinesic" and "acoustic" modes. The visual modes codify several aspects of settings and props as well as lighting and camera movements in the depicted world and in relation to the viewer. The kinesic modes, on the other hand, are dedicated to encode aspects of the on-screen character's non-verbal behaviour and acting including facial expressions, gestures or body movements. The acoustic modes refer to all kinds of sounds, i.e., distinctive qualities of speech, songs, film score, non-verbal noises and other sound effects that create a coherent whole in synergy with all modes (e.g., van Leeuwen, 1999).

A graphical representation in Figure 3.1 shows how the visual, acoustic and kinesic modes are mapped with the resources of the language of film to categorise these elements that contribute to the construal of ironic meaning on screen in a multimodal way.

Figure 3.1 A graphical representation of the visual, acoustic and kinesic modes mapped with the resources pertaining to the language of film
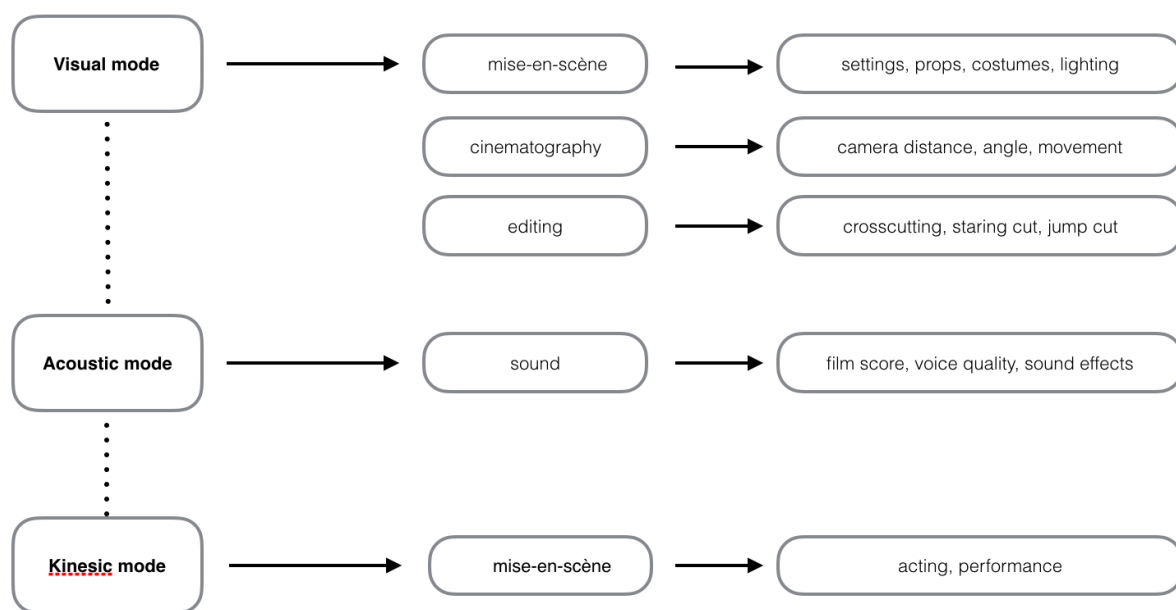
Figure 3.1 represents a graphical representation of the non-verbal modes embedded in the film text. It shows how the visual, kinetic and acoustic modes convey meaning construed by resources pertaining to *mise-en-scène* (e.g., props, gestures)*,* cinematography (e.g., camera movements), editing (e.g., crosscutting, jump cut) and sounds (e.g., film score or character's voice quality). In the next subsection, it has been illustrated how the visual, kinesic and acoustic modes contribute to the construal of multimodal irony on screen.

3.4.2 Visual modes in the construal of multimodal irony

The visual modes integrate resources pertaining to mise-en-scène, cinematography and editing. While mise-en-scène is a result of the combined efforts of art directors, set and costume designers encompassing "all the details that make up the image" such as settings, props, lighting, costumes (Philips, 2000: 35; Bordwell and Thompson, 2007), cinematography determines how resources of mise-en-scène are seen by the audience from various camera angles (Dix, 2008). Editing*,* on the other hand, primarily determines the ordering and the frequency of these camera movements in which viewers can see certain objects on screen (Philips, 2000).

Within mise-en-scène, props and costumes can be ancillary in the construal of ironic meaning through different features of scene decorations and objects. For instance, settings (e.g., landscape) may contribute to construe ironic intent through the geographical location of on-screen characters. Boggs and Pietrie (2008: 82) termed it as "irony of settings" to

indicate an event taking place in a setting that is the opposite of the settings the viewers are expected to see for such an event, for example, a murder in a peaceful setting. Props (e.g., a notebook, a gun or clothing), on the other hand, may indicate not only the socio-economic and occupational status or cultural background, but also information regarding the character's emotional status or attitude (Dix 2008). For instance, in *Sherlock Holmes* (2009) Sherlock's hat, black coat, dark sunglasses with his revolver and pipe suggest he works as a detective during the Victorian period. Concurrently, these objects in the combination of his sloppy clothing and haircut indicates Sherlock's dissociative and disrespectful attitude.

Another resource that brings mise-en-scène into life is lighting – rarely noticeable but equally significant in films (Wharton and Grant, 2005: 44). Dix (2008) distinguishes between high-key and low-key lighting in order to contribute to the construal of the overall meaning in a film scene. For example, optimism and transparency are indicated by low contrast between dark and bright areas on screen (high-key lighting), while anxiety, nostalgia or even terror are signaled through the high contrast between dark and bright areas (low-key lighting). Low-key lighting is also used to highlight the ironic content. It is visualised in Tim Burton's films featuring the use of irony in the dark, moody and gloomy light. Similarly, in Ritchie's Sherlock Holmes films, irony is accentuated by the contrast between dark and bright areas of Victorian London.

Within cinematography, there are three main camera functions, i.e., distance, angle and movement that can support the construal of multimodal irony in the film text. For example, the distance stimulates the position of camera, and thus controls the space between the audience and the depicted world of an image (Baldry and Thibault, 2006). Some of them can be particularly useful when construing ironic meaning. Dix (2008) distinguishes a wide array of shots of different ranges, i.e., extreme long shots in which a human figure is barely perceivable. In long shots, the setting is still prevalent, yet viewers are able to recognise the identity of an on-screen character. The figure becomes clearly discernible from the knees upwards in medium long shots. This type of shots presents the full posture of human body, which can be helpful when illustrating the on-screen character's attitude of dissociation. Medium close-up shots outline a film figure from the waist up. As a result, viewers are able to observe a more detailed picture not only of the characters' posture but also the expression of their emotions. On the other hand, close-ups (or close shots) expose the figure's head and shoulders, while extreme close-ups (or very close shots) locate the film camera predominantly on the character's head and represent the nearest visual distance construed by the camera in which features of the character's face become the focus, in particular the eyes

and mouth (Dix, 2008). A close-up and extreme close-up can be instrumental to visualise facial non-verbal markers of irony such as wide-open eyes, raised eyebrows or a smirk. These non-verbal elements can be illustrated with Example 3.1 (see page 58), when Sherlock is talking to Dr. Watson in a carriage. In this scene a close-up and extreme close-up show Sherlock's facial movements, eyes, gaze, eyebrows, mouth or head movements, thus reflecting his ironic intentions towards Dr. Watson.

Camera angle represents another way of stimulating the spectators' film perception and supporting the construal of ironic meaning. There are three main options in terms of camera angle, i.e., high angle, straight-on angle and low angle (Dix, 2008). While the straight-on angle presents the on-screen characters from a natural position, the two remaining options contribute to the construal of specific meaning. Low angle shots can be particularly conducive to the production of multimodal irony, as it is taken when the on-screen character looks down at the depicted world, and thereby signals dominance, superiority and dissociative attitude towards another subject. High angle shots, however, present the on-screen character above eye-line which results with the audience looking down at the subject indicating weakness and powerless attitude of the on-screen character. An illustration of low and high angle shots in the instance featuring the use of multimodal irony presented in the example below.

Example 3.2

Sherlock and Watson enter the cathedral to save an innocent girl's life that the villain Lord Blackwood intends to kill. Sherlock and Watson act quickly and manage to catch Blackwood and kill his troops before Inspector Lestrade arrives with the police officers. Once Lord Blackwood is caught, Inspector Lestrade walks in and Sherlock says: "Impeccable timing, Lestrade!" with a blank face and wide-open eyes.

Scene 1   Scene 2

As shown in Example 3.2, Sherlock is filmed from a low angle in Scene 1 to demonstrate his dominance and dissociative attitude over Lestrade and his troops while delivering the ironic statement. Inspector Lestrade, on the other hand, is filmed from a high angle in scene

2 demonstrating his position as powerless towards Sherlock and the audience contributing to the construal of multimodal irony in this scene.

Other camera positions are realised through the movements of the camera itself or one of its components. The camera moves along its horizontal or vertical axis. To take horizontal shots, the camera proceeds along the laid track (a tracking shot) or on a wheeled apparatus called dolly (a dolly shot). These types of shots are particularly useful to enhance suspense in thrillers or dramas (Dix, 2008: 28). On the other hand, vertical shots, known as crane shots, are taken above ground level through the camera's falling and raising movements.

Last but not least, editing is another resource that complements cinematography in creating suspense, drama and multimodal irony by providing and withholding information – thus heightening "our awareness of a character's comparative blindness" (Macdowell, 2016: 117). Editing creates links between the sequences of a film into a coherent and narrative story (macro function). Within a sequence, it also links shots according to spatial and timely relations in order to influence the viewers' perception of the film (micro function) (Philips, 2000). There are different editing strategies (e.g., fade, straight cut, the jump cut etc.) that produce different effects. As Macdowell (2016) observes, "it is likely to be true that editing strategies of all kinds may be put in service of ironic presence" (2016: 116). Nevertheless, crosscutting (or parallel cutting)[15] and intercutting[16] may be the most reliable strategies to show the audience "aspects of the fictional world to highlight an ironic state of affairs or potentially reveal a character's limited perspective" (Macdowell, 2016: 117). All these techniques have long been employed to create the effects of dramatic irony, for instance.

3.4.3 Kinesic modes in the construal of multimodal irony

The kinesic modes incorporate a wide array of non-verbal filmic resources pertaining to mise-en-scène such as acting or performance which are integral components of the on-screen characters' bodily actions including facial expressions, gesture and movement, to name just a few (Dix, 2008).

---

[15] An editing technique in which the point of view (p.o.v.) alternates from the action at one location to those of another related action. This method is mostly used to build a dynamic tension, e.g., in the chase scene (Breton, A. (2010). *Film Terminology and Other Resources*. Retrieved June 15, 2017, from: http://www.psu.edu/dept/inart10_110/inart10/film.html)

[16] An editing technique used to speed up a scene and compress the time that would slow an action down, when filming action at two different locations to produce a composite scene, e.g., when two people are talking on the phone from two distinct places at the same time. (Breton, A. (2010). *Film Terminology and Other Resources*. Retrieved June 15, 2017, from: http://www.psu.edu/dept/inart10_110/inart10/film.html)

Acting is predominantly naturalistic and realised on the basis of observed human behaviour and a systematised repertoire of gestures, movements, expressions and intonation reflecting "truth [brought] to life" (Dix, 2008). A constellation of different kinesic actions can also indicate the character's emotional and psychological state of mind, from excitement to nervousness and fear, or reveal an attitude towards another character/object, from being polite to the expression of disapproval and mockery (Dix, 2008). For example, Sherlock in *Sherlock Holmes* (2009) reinforces his ironic message through a combination of several kinesic actions like raised eyebrows, rolling eyes, smirking or winking (see Example 3.1).

Some features of the behaviour or personality of well-known actors, known as "stars", also contribute to the meaning-making process on screen (Philips; 2000), as illustrated on the example of Robert Downey Junior in this section. This phenomenon of star's image is further defined by Dyer (1998: 3) as a "structured polysemy, that is, the finite multiplicity of meanings and affects […] and the attempt to structure them so that some meanings and affects are foregrounded and others are masked or displaced."

3.4.4 Acoustic modes in the construal of multimodal irony

Although films were originally silent, sound has become a vital part of the film composition since 1927 along withthe acoustic modes that have enabled the production of meaning through speech, music and sound effects since then (Dix, 2008). Film scholars predominantly differentiate between diegetic (actual) and non-diegetic (commentative) sounds. *Diegetic sounds* pertain to the category of on-screen or off-screen sounds produced by someone or something visible or implied, respectively, in the film story (Dix, 2008; Wharton and Grant, 2005). For instance, a character pulls a gun and the audience can hear the sound of shooting. Non-diegetic sounds, on the other hand, refer to the category of sound that does not stem from the film story itself but from voice-over and musical scoring, for instance (ibid.).

Diegetic sounds are thus highly significant as they represent the character's voice quality. As discussed in Section 3.2, different features of voice quality reveal information about the character's personality, and thus contribute to the meaning making processes. In the case of irony, for instance, other vocal features such as intonation, pitch and intensity may be seen as markers of ironic meaning (e.g., Attardo et al., 2003; Kreuz and Roberts, 1995). As a way of illustration, Sherlock's voice in *Sherlock Holmes* (2009) expresses his ironic attitude through flat intonation, slower tempo and lower pitch, as shown in Example 3.1.

A film score also performs several functions (Philips, 2000; Dix, 2008). Music intensifies sensory stimulation, incites feelings, motifs and attitudes conveying different meanings in the audio-viewers, while stimulating the imagination, creating the atmosphere in a scene and reinforcing the filmic illusion (Philips, 2000; Fonagy and Magdics, 1972). It also "serves as a vehicle of carrying an ironic message" and it is believed to facilitate "the complexities of ironic musical communication" (Turner, 2016: 214). A film score may also generate a sense of irony when it is disconnected from the mood in the film narrative. Indeed, "music can evoke a strong sense of irony when paired with an image track indifferent to, or in opposition with, commonly accepted emotional connotations" (Turner, 2016: 177). "Musical irony" (Turner, 2016) is illustrated in the film score that Hans Zimmer composed for *Sherlock Holmes* (2009) through the use of rather unconventional instruments like a banjo, cimbalom, squeaky violin, an accordion, dumbeks, a "broken pub piano" or the experibass to produce a kind of quirky music to intensify humorous, dramatic and finally ironic effects. This score is particularly evident in the scene of *Sherlock Holmes* (2009) when Lord Coward tries to find and shoot Sherlock, who is hiding in a stream of smoke. It thus mocks Lord Coward's abilities to catch him. Although Lord Coward is trying to kill Holmes, Zimmer's music intensifies the ridiculousness of this situation, alongside Sherlock's ironic comment at the end of the scene: "There is not time to wait, is there?"

Similarly, sound effects can also intensify emotions, atmosphere or suspense. For example, the sound of a shooting gun in a war film can signal imminent danger. In other scenes, sound effects like seagull cries can determine geographical settings (Wharton and Grant, 2005). Sound effects can also achieve a symbolic dimension (Desilla, 2009). For example, the sound of a heart beating can symbolise danger or nervousness, but also a feeling of sexual excitement.

This section has outlined how different resources of the language of film intersect to relay ironic meaning on screen in a multimodal way through the visual, kinesic and acoustic modes. Film directors have a set of resources at their disposal that can be concurrently employed to portray the speaker's attitude of dissociation and allow the audience to retrieve the ironic meaning behind it. Although these non-verbal resources are no doubt important in the construal of multimodal irony, language plays a pivotal role in its production and retrieval. The next section focuses on the contribution of narration to the process of construing ironic meaning in the film text, highlighting the principal functions of film dialogue.

3.5 Verbal Modes in the Construal of Multimodal Irony

Beyond the non-verbal sign systems, the verbal modes are (in the spoken and written form) an indispensable element of the film text which allows the viewers to "understand the visuals, repeats their information of emphasis, interprets what is shown and explains what cannot be communicated visually" (Kozloff, 2000: 39). Narration also helps filmmakers convey their own ideas and film story to the audience. In order to uncover the intentions of film directors and guide the spectators through the interpretation of the film plot, film dialogue has to fulfil two fundamental functions: a narrative function and an aesthetic function (Kozloff, 2000). More importantly, in the context of this thesis, it is significant to explore how these functions participate in the production of multimodal irony in the film text.

For instance, the aim of the narrative function is to arrange events, causality, characters, emotions and information into the film story (Kozloff, 2000; Philips, 2000). That is to say, dialogue creates the fictional world of narration helping the audience to determine when, where and how an action takes place through the identification of cause-effect relations (Bordwell and Thomson, 1997; Kozloff, 2000; Desilla, 2009). Analogously, irony in film dialogue is arranged in a certain chronological order and unfolds through a series of events linked by causality in a given physical space. In this way, it could be argued that irony contributes to the plot development managing the viewers' access to information. This is evident in "omniscient narration" (Bordwell, 1985) in which a narrator has more information available than any other character. Irony thus occurs in the narration in which a narrator knows more than other on-screen figures and makes this fact clear to the audience (Tan, 1995). In dramatic irony, on the other hand, information is shared with the audience but the on-screen characters do not notice it. In this case, the character understands the literal meaning of something while the audience perceives its ironic meaning (Beaver, 2007; Boggs and Petrie, 2008).

Another narrative function pertains to the character's revelation in the diegetic world. Dyer (1998: 106) states that "a character is a construct from the very many different signs deployed by a film". These signs may play an important role in expressing the character's personality, social class or emotional condition as "each time a character opens his mouth, filmgoers learn more about him" (Kozloff, 2000: 43). However, not only the character's appearance but also film dialogue can reveal information about the protagonists and yield a deeper insight into their psychological profile (Kozloff, 2000). A good case in point is Sherlock in *Sherlock Holmes* (2009) whose cutting remarks and ironic comments towards

Inspector Lestrade indicate his mocking attitude, and thus the ironic nature of his personality. Additionally, dialogue can also portray a specific type of relationship between two protagonists. For instance, Sherlock and Dr. Watson in *Sherlock Holmes* (2009) argue like "an old married couple" (Murray, 2009) but in reality, the wordplay (Lyall 2009) and the banter (Scott 2009) between them signal the special ironic humorous character of their "bromance".

The aesthetic function, on the other hand, refers to the artistic nature and resources of language (Kozloff, 2000). In other words, it provides the spectators with aesthetic pleasure and intellectual stimulation (Philips, 2000), as they desire "more emotions, more spectacle, more stimulation of the emotions and intellect" (Philips, 2000: 25). One of the means at filmmakers' disposal is verbal interaction and the use of "a more sophisticated talk-humour" in film dialogues (Hayward, 2000: 90). Scriptwriters resort to different aesthetic effects and linguistic tools through the use of poetic effects, humour, allegory, metaphor, and more importantly irony (Kozloff, 2000). As Boggs and Petrie (2008: 80) observe, "irony adds an intellectual dimension and achieves both comic and tragic effects at the same time".

As mentioned above, many filmmakers use dialogue to convey jokes and humour to change the mood or relax the viewers before the next scene in action or thriller films. Dialogue has thus become "the comic engine of the text", particularly in sound comedies (Kozloff, 2000: 54). Garrand (1984: 243) distinguishes several subcategories of humorous dialogue including "understatements", and "sarcasm", among others, which can be considered as different types of verbal irony. As discussed in Section 3.2, while irony can be successfully conveyed through the visual and acoustic modes, "language greatly expands film's ironic capabilities" (Kozloff, 2000: 54). Similarly, to other figures of speech, irony can also be regarded as a manifestation of "linguistic indirectness" and its effect depends on the recognition of ironic meaning (Desilla, 2009: 118). In metaphor likewise, ironic meaning is also hidden "beneath the surface awaiting to be discovered by the audience" (Desilla, 2009: 119). Irony has often been construed through "repartee" which Abrams (1993: 220) defines as "a contest of wit in which each person tries to cap the remark of the other, or to turn it to his or her own advantage". In other words, the protagonists using ironic repartee can express their attitude of disapproval through banter, cut wit as well as deliberately fabricated meaning and play with each other's utterances to "mock or disparage the people or situations being referred to" (Gibbs and Colston, 2007: 588).

Humorous irony also is founded on the notion of incongruity and superiority (Vandaele, 1999/2002). Incongruity arises when a linguistic expression, for instance in the

film dialogue, contradicts the expectations in a particular context, and hence creates humour. In addition, incongruity not only adds value to the humorous functions of irony but it also construes ironic meaning through ironic riposte and ironic metaphor. As Nilsson comments (2013: 113) below:

> In the film which is dominated by its dialogue the characters, for instance, use the ironic riposte and the ironic metaphor in dialogue sequences designed to signal incongruities and cue a reversal of meaning.

Superiority, on the other hand, arises when the viewers have been given access to the information that other characters do not have, and as a result can identify the incongruity and feel superior to the character being mocked (Vandaele, 19999/2002; Desilla, 2009).

3.5.1 Pastiche, intertextuality and irony

Verbal and non-verbal resources can express ironic meaning in unison, referring viewers to other resources or earlier parts of the film. This practice of referring to previous works has been termed as "pastiche" (Dyer, 2007). This widely and loosely used term originates from the Italian noun pasticcio. In the early Renaissance, pasticcio described "a genre of painting of questionable quality […] that synthesised - "stirred together" styles of major artists, often with seemingly fraudulent intentions, i.e., to deceive viewers and patrons" (Hoesterey, 2001: 1).

In the late 17th century, this Italian concept was further developed in France and became known as *pastiche* and referred to the practice of reusing pieces of music, art or written works in new productions and entities (Hoesterey, 2001: 8) to celebrate great works in a new light without copying them, as pastiche imitates the universal features of the arts to which it refers, rather than precisely reproducing them (Dyer, 2007).

Pastiche can be thus defined as an evaluative open imitation shaping a particular kind of reference to other authors and genres as well as to visual arts, paintings, music or written works. It aims to imitate style, technique and motifs of previous works or periods in art, literature, and more recently in films (e.g., Hoesterey, 2001) fostering creativity and celebration of the great works of the past that are illustrated again in shows, movies or books. It does so from a novel perspective through the mixture of genres and styles so that viewers can celebrate the presence of past events, pieces of arts or music. An example of pastiche in Ritchie's Sherlock Holmes films (2009, 2011) is visible when he incorporates novel trademarks of Holmes from other films to present Sherlock in a new light. In this case, this involves combining the traditional representation of Sherlock Holmes with more violent

characters like James Bond and a witty and ironic characters like Dr. House. As a result, pastiche intends to introduce humorous elements into a new piece of work without the intention to parody or mock it (e.g., Hoesterey, 2001).

As the notion of pastiche encompasses a wide array of subcategories, in this thesis, its interpretation will focus on "the study of reference, allusion and intertextuality" with a particular emphasis on the concept of intertextuality, which is particularly useful in the creation of ironic meaning on screen (Dyer, 2007: 22). Originally, the notion of intertextuality was based on the premise that "(all) text is always constituted out of elements from other texts" (Kress, 2000: 139). Intertextuality in film refers to relationships with other films or establishes links with an earlier period in film history through the combination of the verbal, visual, kinesic and acoustic modes (Hayward, 2000; Philips, 2000; Stafford, 2007). This may be performed by quoting a line of dialogue or through a series of shots, camera movements or music score (Philips, 2000). Intertextuality can be realised through dialogue, cinematography, editing, music or mise-en-scène, or through a co-deployment of all of these modes. The intertextual film connections do not aim to convey fixed meanings but "allow the spectator to […] engage with the film in a playful way" (Stafford, 2007: 83-84).

Intertextuality can also serve as a useful tool to convey ironic meaning in films. In his study, Zabalbeascoa (2003) analysed verbal irony in *Transpotting* (Danny Boyle, 1996) and identified phrases, idioms and ironic utterances that refer to a certain type of text, such as recipes. That is, a clear intertextual link with the language of cookery recipes is presented in the way Renton tells the audience how he is going to overcome his addiction. In the case of *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011), an intertextual relationship can be found with other films or series, such as *House, M.D.*, in which researchers have argued that irony also plays a pivotal role (e.g., Dynel, 2014). As it is in the case of Gregory House, Sherlock Holmes also often expresses irony both verbally and non-verbally. Like House, Holmes has no mercy for other on-screen characters, and expresses his cutting remarks towards them in an open manner. Both characters express their attitude of dissociation through the visual, kinesic and acoustic modes, such as a blank face, wide-open eyes, smirking or through the use of flat intonation and low pitch.

While quick-witted dialogues, verbal play and, more importantly, irony heavily dominated the American genre of screwball comedies in the 1930s and 1940s (Kozloff, 2000), contemporary cinema extended the use of ironic dialogues to novel film genres at the end of twentieth century (Philips, 2000). Filmmakers have begun to "play" with the audience

incorporating irony and wit into action movies (Philips, 2000). This development of cinematic traditions thus resulted in a mixing of genres heralding the advent of a new type of films. These "hybrid" films merge the elements of different genres and intertextual features (Philips, 2000). For example, *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011) combine irony with the elements of comedy and action movies through numerous intertextual relations.

Sherlock Holmes films embody several features related to the genre of screwball comedies. To start with, there is a discrepancy between literal and ironic meanings which is a characteristic feature of screwball comedies and plays an important narrative role in dialogues between Sherlock and Dr. Watson (see Example 3.1). In addition, similarly to screwballs, Sherlock Holmes movies also represent "the battle of the sexes" (Hayward, 2000: 89). That is to say, through wordplay, undercutting and sexual innuendos, the male and female protagonists of screwball comedies are often too proud and independent to descend into sloppy romanticism and use ironic remarks as a common speech act to subvert the language of love. This can be observed in the example below, which features the relationship between Sherlock Holmes and Irene Adler, who often interlace their dialogues with ironic remarks, inviting the audience to partake in their love game:

Example 3.3

> Irene pays a visit to Sherlock in his apartment, while he is still asleep. Sherlock wakes up immediately and both start a conversation fulfilled with cutting comments. Sherlock's widely open eyes, raised eye brows, smirks and head movements also help him express his ironic intensions. Irene looks at the documentation left on his desk and says to Sherlock: "I see you are between jobs.' In a medium close-up shot, Sherlock looks at Irene from a high angle and responds: "And you are between husbands?"

This love-hate relationship between Sherlock and Irene is not sufficient reason to classify the Ritchie's two productions (2009, 2011) as pure comedies. As both films combine ironic dialogues with gunfire, explosions or fight scenes, they appear to represent a more "action-oriented Sherlock Holmes" with wide-ranging irony (Murray, 2009). The type of film that involves spectacular and fast scenes with briskly action-packed narratives is called "action film". Action movies combine aspects of science-fiction, spy thrillers, fantasy and martial art along with crime and war films (e.g., Hayward, 2000). The main idea behind action movies is that the spectators lose a sense of reality and direction in a flurry of sound and images, which enhances the level of suspense and engagement in the action (e.g., Hayward, 2000).

Through a number of intertextual relations, Sherlock Holmes films (2009, 2011) link features of screwball comedies with a wide array of action-packed scenes, thereby constructing a pastiche of several genres. For example, in Ritchie's story-lines, Sherlock resembles the fighting Bruce Lee in *The Big Boss* (1971) through the use of "the martial art baritsu" (Lyall, 2009). In regard to spy thrillers and crime films, the intertextual relationship between Sherlock Holmes and James Bond is clearly discernible, not only through spectacular action encompassing explosions, bombs and a series of violent shoot-outs, but also through verbal interaction, largely involving ironic and cutting dialogues. On the other hand, the link to science-fiction or fantasy movies in *Sherlock Holmes* (2009) is found in the use of black magic practices that connect villain, Blackwood, with the black magician Lord Voldemort in Harry Potter films. Such magic has been termed as "cosmic irony" presenting life as a game of absurdity in the presence of a supreme being or creator (Boggs and Petrie, 2008: 84). The above genre analysis indicates that Sherlock Holmes films (2009, 2011) seem to be a hybrid combination of different film genres, integrating elements of comedies and action films. These movies can be thus considered "a snappy, fast-paced, albeit confusing, action comedy" (Murray, 2009) featuring the use of irony.

3.6 A Definition of Multimodal Irony

As explained in Section 3.2 above, non-verbal aspects of irony contribute to the construal and retrieval of ironic meaning in face-to-face and prefabricated communication acts, facilitating the generation and interpretation process of verbal irony. Film communication, however, seems to add a further layer of complexity.

The discussion in Sections 3.3, 3.4 and 3.5 highlighted that the film text is a polysemiotic medium in which "movements, sounds, gestures, objects, [and] colours are as important as verbal utterances and their meaning cannot be ignored" (Chaume, 1997: 315). On this account, filmic meaning is produced through the combination of non-verbal modes, realised through the language of film, in tandem with film dialogue.

As illustrated in Section 3.4, ironic meaning can be conveyed through the amalgamation of the visual, kinesic and acoustic modes with dialogues in film narration. Therefore, when analysing ironic meaning in the polysemiotic context of films, the notion of irony should be re-defined. A new definition of "multimodal irony" has been proposed, drawing on insights from Sperber and Wilson's operationalisation of irony and extending them for application in a multimodal context.

In the light of their echoic theory, multimodal irony in film can be defined as a double-layered phenomenon communicated by filmmakers through the on-screen characters' dissociative attitude towards the proposition echoed and expressed through the combination of the visual, acoustic, kinesic modes, in unison with dialogues. That is, multimodal irony can be conveyed on two layers: vertically, i.e., between the on-screen characters, and horizontally, i.e., between the on-screen characters and viewers of audiovisual productions (Vanoye, 1985). The latter gives the spectators additional possibilities to retrieve ironic meaning from a visual mode, possibilities which are not available to the on-screen characters. Hence, the attitude of dissociation expressed by the on-screen protagonist towards the proposition echoed can be targeted at another on-screen character or directly at the viewers.

Furthermore, viewers can supplement on-screen details with off-screen knowledge to process the intended ironic meaning, particularly when irony conveys culture-specific meanings. For instance, viewers can use relevant background information, such as the on-screen characters' prior utterances, as well as contextual clues, sociological, historical, cultural and encyclopedic knowledge and personal experience. More importantly, the spectators can take advantage of the fact a film is a polysemiotic medium to retrieve or enhance their understanding of the intended ironic meaning conveyed non-verbally by processing verbal and non-verbal modes simultaneously to create a meaningful whole.

The choice of modes that filmmakers use to convey ironic meaning is predominantly determined by the intended effect. For example, with a wide array of non-verbal resources like props, camera movements or kinetic behaviour in action comedies, filmmakers enhance entertainment and stimulate intellectual desire, thus encouraging viewers to partake in the creation of meaning (Kozloff, 2000; Desilla, 2009).

3.7 Conclusions

This chapter focused on the study of the predominantly non-verbal nature of irony in the film text. A critical overview of acoustic and kinesic markers was presented, showing that they can aid in the generation and interpretation of ironic utterances (e.g., Rockwell 2000; Attardo et al. 2003; Gibbs 2011) highlighting the complexity of ironic meaning as an object of cultural expression. Subsequently, drawing on the multimodal approach, this chapter discussed how the visual, kinesic and acoustic modes, in unison with verbal dialogues, contribute to the construal of multimodal irony in the film.

This discussion revealed that irony "represents a phenomenon that is more nonverbal than verbal in nature" (Rockwell, 2000: 484). In other words, both voice modulations like flat intonation or lower pitch (e.g., Kreuz and Roberts, 1995; Antolli et al., 2000) and kinesic behaviour such as "a blank face" or raised eyebrows (Attardo et al., 2003) can be instrumental in the generation and interpretation of ironic utterances, particularly when the speaker and the listener do not share the same contextual background (Yus, 2000).

Although these findings yield interesting insights in themselves, a number of shortcomings of existing approaches were identified in the analyses in Section 3.2, regarding the present study of multimodal irony. First and foremost, studies on the notion of irony have tended to explore how the individual non-verbal elements contribute to the construal and comprehension of ironic meaning rather than to explain how these elements function simultaneously.

Another point that has largely been neglected so far is the existence of cultural differences in the interpretation of ironic intent. While several researchers (e.g., Rockwell, 2000; Antolli et al., 2000) have investigated how multimodal irony is generated in various languages, the way in which audiences, particularly the Polish ones, retrieve ironic meaning in the foreign culture and language of the film text remains largely unexplored.

Subsequently, drawing on Baldry and Thibault's (2006) approach to multimodality, visual, acoustic and kinesic modes have been defined. These support the construal of multimodal irony on screen through a combination of resources pertaining to the language of film, such as the soundtrack, mise-en-scène, cinematography and editing. These non-verbal film resources have been shown to aid in the construal of irony in the film text, and hence they contribute to the comprehension of ironic meaning by the audience. Film directors have a whole array of resources at their disposal, such as camera angle, settings, editing techniques or music score to produce intended effects on screen, by manipulating and stimulating the way in which meaning and, in particular, the way in which ironic meaning is perceived and understood by the viewers.

Following these lines of reasoning, Section 3.5 has also revealed that verbal and non-verbal modes cannot work in isolation as film dialogue, as they produce a meaningful whole. This cooperation between verbal and non-verbal modes in the film text contributes to the multimodal construal of irony on screen, not only in comedies, but also in action comedies, the genre to which the films *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011) belongs. Additionally, the meaning-making process inherent in the

construal of ironic intent in film communication can be supplemented through a film's intertextual connections with other films.

As discussed in this chapter, ironic meaning can be construed in a multimodal way through the combination of verbal, non-verbal and intertextual resources in the film itself. It is clear then that alternative approaches to the study of irony must be developed that take account of this complex and multi-pronged relationship. The definition of multimodal irony has therefore been adapted to this multimodal environment.

Nevertheless, far too little attention has been paid to the way in which multimodal irony affects users of multimodal texts. It is thus necessary to explore how viewers experience and understand films, and to establish which resources they use to retrieve meaning in the film. On the subject of irony, it is also essential to explore the extent to which non-verbal resources contribute to the multimodal construal of irony in audiovisual productions. The next chapter outlines the methodology used in this study to analyse the construal and reception of multimodal irony in the specific film being studied.

4 METHODOLOGY

4.1 Introduction

The previous chapter outlined the theoretical insights from established studies in multimodality and film studies and explored how meaning, and more importantly ironic meaning, can be construed in the multimodal texts like films. This chapter offers a detailed description of the methodological apparatus which has been designed to examine the extent to which the participants are able to retrieve the intended ironic meaning that is composed and rendered in the films under analysis.

This chapter is structured as follows. Section 4.2 presents the benefits of triangulation in the context of the current research project. In Section 4.3, the criteria for selecting participants are set, while in Section 4.4 the rationale underlying the choice of the film texts to be used in the current study is provided. Subsequently, Section 4.5 presents a descriptive component, namely, an adapted model of multimodal transcription used to examine the construal of ironic meaning in the film text under analysis. In Section 4.6 the interactionist component is outlined, that is, the questionnaire with coding apparatus to assess whether and/or to what extent respondents were able to retrieve multimodal irony in the selected subtitled and voiced-over clips. Section 4.7 provides an overview of the experimental component featuring the use of the eye-tracking technology. It is examined how eye tracking can help me understand viewers' visual behaviour while watching the excerpt in the present study. This chapter concludes with Section 4.8, which discusses and summarises the application of triangulation of methods to the study of multimodal irony on screen.

4.2 Triangulation of Methods for the Present Study

Triangulation is a combination of at least two research methods to collect and examine data in the same study. It involves crosschecking and verification of findings that have been obtained from various sources. This process is meant to produce data that are less biased by a researcher (O'Brien and Saldanha, 2014). Thus, the same phenomenon or variable should be studied with more methods than one to receive valid findings. The usefulness of this approach has been confirmed in a number of audience reception studies examining the response to the AVT content so far (e.g., Caffrey, 2009; Kruger et al., 2013; Orrego-Carmona, 2014b; 2015; Szarkowska et al., 2016) and on this premise it has been applied to the current thesis.

In the context of the present study, the triangulation of methods has been considered particularly useful combining multimodal transcription (see Section 4.5), questionnaire (see Section 4.6) and eye tracking (see Section 4.7). While the apparatus of multimodal transcription enables me to conduct a detailed and meticulous study of the construal of multimodal irony on screen, it can lead to a possibly biased judgement. To verify whether viewers actually looked at these non-verbal elements that are indicated in the multimodal analysis, the eye-tracking technology has been incorporated into the current design. Eye tracking delivers objective data of the gaze logging as to how viewers consume on-screen information with subtitles and voice-over. Although eye-tracking provides researchers with valuable insights, it cannot measure what the participants were thinking about when watching the instances of multimodal irony in the sequences under analysis. In order to enhance confidence and credibility of the present research account, the eye-tracking data has been complemented with a questionnaire as a form of elicitation of conscious answers to structured questions from the respondents participating in the eye-tracking study.

4.3 Criteria for the Selection of Participants

In the present study, the inclusion of different language groups is particularly important in the context of the Polish audiovisual scene due to economic, social and historic reasons, as discussed in Sections 1.2 and 1.3. It is thus vital to explore how viewers at different levels of English-language competence consume both subtitled and voiced-over videos featuring the use of multimodal irony in order to establish which form of audiovisual translation is more suitable for a certain group of viewers depending on their level of English proficiency.

Therefore, the main criteria for selecting participants was a level of English proficiency. In order to bring out differences and similarities, subjects with a low, medium and high level of English were included in the experiment. The inclusion of different language abilities is crucial as "participants engage differently with the content depending on their language skills" (Orrego-Carmona, 2014b: 88), for instance, with the subtitled products (Orrego-Carmona, 2014b; 2015). Although one can assume that viewers with a high level of English proficiency may not need subtitles or voice-over to understand the film text in the source language, it is still highly important to explore to what extent this assumption is legitimate in regard to the audience participating in the current experiment. It has already been shown that, for instance, Spanish viewers with a high level of English still rely on subtitles to a certain extent (Orrego-Carmona, 2014b; 2015).

In addition, the general proficiency of English language among university students and graduates in Poland was reported as relatively high, as several opinion poll agencies reveal. For instance, the most recent study[17] (2013) shows that out of a population of 58 000 students, 22% respondents reported to have a very good command of English at C1 level, while 26% and 27% of those surveyed reported to be at B2 and B1 level, respectively. Only 9% and 3% of students revealed that their English knowledge is rather basic, that is, at A2 and A1 level, respectively. This information is vital considering the fact that English was introduced to the Polish education system as a foreign language after the fall of communist regime in 1990s, which appears as a relatively recent movement in comparison to other Western European countries (Karoń). In the current project, the participants completed a self-assessment questionnaire (see Appendix 2) rating their listening comprehension skills in English, and subsequently were allocated to one of the three groups representing high (C1, C2), medium (B2, B1) or low (A1, A2) level of proficiency in English that have been mapped on the Common European Framework of Languages[18] (CEFL, see Appendix 1). The participants with a high, medium and low level of English have been labelled for the following abbreviations throughout the whole thesis, that it, HLPs, MLPs and LLPs, respectively.

Other set of criteria involved nationality, age and familiarity with the source material under analysis in order to exclude additional variables. As a result, all participants were native Polish speakers and aged between 18-35. The age group 18 – 35 years is also vital in the light of this study since these members of Polish population represent the two most common groups (18-24 and 25-35 years old) that visit cinemas and watch films online the most frequently, as Tomo Group[19] reported. As a result, these individuals were recruited to the eye-tracking experiment who have been frequently exposed to the multimodal content. Special care was also taken to select the respondents who knew the character of Sherlock Holmes and watched at least one of the two films SH1 and/or SH2 to ensure that all participants are on a similar level of understanding of who Sherlock Holmes was and how he was presented in Guy Richie's films (2009, 2011). Additionally, all participants had normal vision or corrected-to-normal vision (by wearing glasses or contact lenses). The

---

[17] The most recent study regarding the levels of English proficiency among Polish students and university graduate was published by Opinion Poll Agency in Poland and it is available at: http://bip.um.szczecin.pl/UMSzczecinFiles/file/WOIiB/Broszura_PL.pdf

[18] Common European Framework of Languages is a guideline describing levels of achievements of learners of foreign languages across Europe and in other countries available at: https://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf

[19] Opinion Poll Agency in Poland regarding visits in cinemas and watching films online http://www.e-polskiekino.pl/Raport2.pdf

actual purpose of the experiment remained confidential. The participants were only informed about the research goals in very general terms. That is, they were explained that the goal of the present study was to enhance their spectatorial experience of watching foreign films translated with subtitles and voice-over.

4.4 Rationale for the Selection of Data

Two filmic adaptations of famous novels by Sir Arthur Conan Doyle: *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011) directed by Guy Ritchie were selected as my data set in the current project. These films have been labelled as SH1 and SH2, respectively, throughout the thesis. The procedure for selecting six scenes featuring the use of irony in the subtitled and voiced-over versions of SH1 and SH2 has been described in Subsection 4.5.3.

The rationale underpinning the choice of the two films is primarily driven by the presence of multimodal irony conveyed through a combination of the verbal and non-verbal visual, acoustic and kinesic modes in the communication between Sherlock Holmes and his companion Dr. Watson in which Robert Downey Jr. and Jude Law play Holmes and Watson, respectively. As shown in the films, both actors exhibit an ironic attitude in the rapport between the two on-screen characters with a plethora of mockery. This Holmes-Watson relationship has been best described as "the relish of language and the cerebral tennis matches that go on between them as they unravel this mystery" (Lyall, 2009). The ironic and humorous character of the relationship between Downey's Holmes and Law's Watson has also been noticed by reviewers saying that "[b]est of all is the banter between Mr. Downey and Mr. Law" (Scott, 2009).

Apart from banter and mocking jokes, Robert Downey Jr. also expresses the attitude of dissociation through several non-verbal resources that largely contribute to the construal of certain aspects of personality traits and behaviour of the main character - Sherlock Holmes. This "louche, slightly wicked-looking character" (Lyall, 2009) thus relays irony through his humorous costumes, eccentric behaviour, facial expression and gestures. Downey's portrayal of Holmes has also been appreciated by Burr (2011) who refers to "his characteristic twitchy wit and haggard insouciance". Interestingly, Downey, who seems to be a perfect embodiment of Holmes' ironic personality, is also very ironic himself. As Burr (ibid.) noticed, "irony is a gift Downey likes to share with audiences". Moreover, Sherlock's smirks, cheeky laugh, disapproving look, flat intonation and emotionless face expression underscore his ironic attitude also towards other on-screen characters in the film like

Inspector Lestrade, Irene Adler, Lord Coward, Lord Blackwood or Professor Moriarty. Another important reason for selecting these filmic versions of Sir Arthur Conan Doyle's detective stories is the fact that both SH1 and SH2 have been very successful worldwide, including the UK and Poland. For instance, *Sherlock Holmes* (2009) was widely acclaimed and placed 4th in the worldwide rankings in 2009, while *Sherlock Holmes: A Game of Shadows* (2011) was placed 6th on the global film market in 2011, as the Statistical Yearbook reports. Both productions were ranked 8th in British cinemas in 2009 and 2011, while in Poland *Sherlock Holmes* (2009) was listed as the 21st most popular film in cinemas. Two years later, *Sherlock Holmes: A Game of Shadows* (2011) was ranked 13th (Yearly Polish Box Office 2010, 2012).

Although Sherlock Holmes films were less popular in Poland, there are a number of reviews of SH1 and SH2 published on Polish Internet portals that present Ritchie's films in a positive light (e.g., Pietrzyk, 2010). Also, Internet-users scored *Sherlock Holmes* (2009) with 7.8 [20] points and *Sherlock Holmes: A Game of Shadows* (2011) with 7.6 [21] points out of 10 according to FILMWEB[22] – an online Polish database devoted to films and cinema — which indicates that both films were well very received by Polish viewers.

Since both films are very popular worldwide as well as in Poland it has been assumed that the participants in the current study will be familiar with the character of Sherlock Holmes played by Robert Downey Jr., to some extent, and this knowledge will support their recognition of a number of the verbal and non-verbal resources construing multimodal irony in SH1 and SH2. Therefore, the two Sherlock Holmes films have been found to be very productive for the analysis of multimodal irony in this thesis, as also shown in the next section on the multimodal concordance.

4.5 Descriptive Component: Multimodal Transcription

This section introduces an adapted model of Baldry and Thibault's (2006) multimodal transcription tool as a proposed methodological device to assist with the selection of the excerpts from SH1 and SH2 for the experimental part, and later with the analysis of the

---

[20] The scoring of Polish Internet-users to the film *Sherlock Holmes* (2009) can be found in the online Polish database of FILMWEB available at: http://www.filmweb.pl/Sherlock.Holmes

[21] The scoring of Polish Internet-users to the film *Sherlock Holmes: A Game of Shadows* (2011) can be found in the online Polish database of FILMWEB available at:
http://www.filmweb.pl/film/Sherlock+Holmes%3A+Gra+cieni-2011-556864

[22] FILMWEB is the largest in Poland and second largest worldwide online database devoted to the publication of rankings and information on films and cinema available at: www.filmweb.pl

descriptive data. Subsequently, it is illustrated how multimodal concordance works in practice on the basis of one of the examples selected from SH1.

4.5.1 Baldry and Thibault's adapted model of multimodal transcription

In "a dynamic complex environment" film dialogue alone is not sufficient to account for the complexity of the multimodal nature of a film (Dyer and Pink, 2015). As shown in Section 3.4, there are the visual, kinesic and acoustic modes that produce meaning, in particular ironic meaning in concert with film dialogue. A methodological tool is now required that allows to effectively gauge the contribution of these elements of film language to the process of multimodal meaning-making in a more rigorous and systematic way.

Multimodal transcription is a well-established and systematised methodology in film analysis that enables the dissection of multimodal texts like films. Although scholars have devised a range of multimodal transcription conventions, the model put forward by Thibault (2000) and Baldry and Thibault (2006) has been shown to lend itself to adaptation for the analysis of the film text with interlingual subtitles (Taylor, 2003, 2004; Desilla, 2009). Therefore, in the context of this study, an adapted version of Baldry and Thibault's (2006) apparatus was employed to select and analyse the clips featuring multimodal irony. Only these non-verbal resources were transcribed that are believed to contribute to the construal of multimodal irony in the subtitled and voiced-over clips.

To conduct a detailed dissection of the chosen film excerpts, Baldry and Thibault (2006) proposed a micro level approach to transcription which involves "a detailed description of the semiotic resources used in the meaning-making process" (Baldry and Thibault's, 2006: 166). Since this thesis analyses the non-verbal semiotic resources at play their model has been adapted to gauge how these semiotic resources operate in unison in the relay of verbal irony in both the subtitled and voiced-over clips.

Multimodal transcription starts with the selection of frames. A frame constitutes a minimum unit of footage and it is recorded 25 times per second in a video file. Due to time and space limitations of this thesis, it is unfeasible to transcribe each frame, as the analysis might be too complex (Norris, 2007). Therefore, in the context of the current study, one frame per second is selected, which contains salient elements of cinematography, mise-en-scène, editing and sound composing ironic meaning on screen (Baldry and Thibault, 2006; Desilla, 2009). The presence of multimodal irony can be established when a number of elements pertaining to the language of film concurrently are co-deployed to produce irony on screen.

The multimodal transcription table displays vertically aligned frames in each row illustrating clips featuring the use of multimodal irony. Each row is flanked by a number of columns; each column illustrates how different verbal and non-verbal resources interact to construe multimodal irony. The adapted apparatus of multimodal transcription comprises seven columns. The first column, **Column 1,** indicates a numerical order of visual frames in each row which is located on the left edge of the table, while **Column 2** specifies the time progression of a frame under scrutiny with accuracy to seconds. **Column 3** displays the selection of visual frames chronologically. **Column 4** contains a list of these meaning-making resources that are believed to construe multimodal irony on screen, as discussed in Section 3.4. These parameters have been presented (based on Baldry and Thibault, 2006; Dix, 2008; Philips, 2000) in Table 4.1:

Table 4.1 List of parameters with abbreviations and explanations used in the multimodal analysis of irony in SH1 and SH2

Parameter	Abbr.	Explanation
Camera position	CP	Stationary: horizontal shots e.g., *dolly*; vertical shots e.g., *crane shots*
		Moving: horizontal shots e.g., *pan;* vertical shots e.g., *tilt shot* and *zooming*
Camera distance	CD	Distance between the depicted word of the image and the audience and includes: VCS (very close shot), CS (close shot), MCS (medium close shot), MLS (medium long shot), LS (long shot) VLS (very long shot)
Camera angle	CA	HA (high angle) SA (straight-on angle) LA (low angle)
Visual collocation	VC	Visual elements that contribute to the construal of meaning such as settings, props and costumes.
Lighting	L	LK (low-key lighting) HK (high-key lighting)
Editing	ED	Different editing techniques (e.g., *continuous*, *fade, straight cut, the jump cut*).
Visual focus	VF	Interactive relation between the on-screen characters and the viewer, e.g., the direction of gaze may indicate ironic intent (Williams, 2009).

Column 5 tracks down meaning-making resources pertaining to the kinesic modes, i.e., body movements, facial expressions or gestures which are deemed highly ancillary in the composition of ironic meaning (e.g., Attardo et al., 2003). Drawing on the studies discussed in Subsection 3.3.2, a set of notational conventions has been proposed to transcribe kinesic features that have the potential to compose multimodal irony in the scenes under analysis, as illustrated in Figure 4.1 below. Each notation of kinesics is labelled with a graphical representation (adapted from Muecke 1978; Attardo et al. 2003; Williams et al. 2009; Wilson and Sperber 2012):

Figure 4.1 Kinesics notation with a corresponding graphical representation

Kinesics	Graphical Representation
Eyebrows	△ (raised)
	▽ (lowered)
Eyes	◎ (wide open)
	◍ (rolling eyes)
	☐ (disapproval)
	↺ (winking)
Facial expressions	◇ (blank face)
	⏢ (wry smile; smirk)
	☺ (smile)
Body movements	� (a resigned shrug)
	⌐ (nodding)
	⊟ (a weary shake of the head)
	⊗ (laugh)

In addition, a brief description of kinesic actions in the scene is provided in either square or round brackets in Column 5. The square brackets are used to indicate kinesic movements that occur either simultaneously or are embedded into each other, while round brackets are used to specify several kinesic actions that occur in a sequential order. **Column 6** distinguishes verbal from non-verbal sounds in the film text. In Figure 4.2, a set of conventions was introduced with a simplified version of Baldry and Thibault's (2006) soundtrack notations. These notations involve acoustic markers of ironic tone of voice that are labelled with a graphical representation (adapted from Schaffner, 1981, 1982; Rockewell, 2000; Antolii et al., 2000; Bryant, 2011; Attardo et al., 2003; see Section 3.3.2)

Figure 4.2 Sound notation with a corresponding graphical representation

Sounds	Graphical Representation
instrumental music	♫
lengthened syllables	(!!)
speech; dialogue (Sherlock, Watson, Irene, Lestrad, Mary, Clarkie)	SH, WAT, IR, LES, M, CL
non-verbal and non- musical sound	☼
no music	[☼ silence]
degree of loudness	**pp** (very soft); **p** (soft); **n** (normal); **f** (loud); **ff** (very loud)
tempo	**S** (slow); **M** (medium); **F** (fast)
pitch	**LP** (lower); **MP** (medium); **HP** (high)
intonation	↘ (falling); ↖ (raising); ↔ (flat)
pause; long pause	#; ##

Furthermore, SI stands for simultaneous when two or more sounds occur at the same time. Finally, the on-screen characters' speech, film narration and thoughts are transcribed in italics to differentiate verbal from non-verbal soundtrack. **Column 7** includes a back-translation (BT) of the Polish subtitled and voiced-over clips under scrutiny due to their salience as a subject of analysis in the present project. A slash (/) is inserted to divide a two-line subtitle, while in the voice-over version a slash (/) signals the division of dialogues between the on-screen characters. Each transcribed scene is allocated an individual code for instance **SH1_1.** In this example, SH indicates a Sherlock Holmes film, the number 1 following the abbreviation of SH indicates the first of out the two Sherlock Holmes films, i.e., *Sherlock Holmes* (2009). Analogously, SH2 indicates the second of the two Sherlock Holmes films, i.e., *Sherlock Holmes: A Game of Shadows* (2011). The number given after the underscore (_) determines the chronological order in which each clip is displayed in regard to the whole film text under scrutiny (SH1 and SH2). Additionally, the abbreviations SUB and VO were used to refer to subtitles and voice-over, respectively.

In order to show how the outlined above model of multimodal transcription has been employed in practice, two examples from SH1 have been presented illustrating how verbal and non-verbal resources are combined together to compose multimodal irony. In Table 4.2

the selected frames were transcribed to see how multimodal irony is construed and relayed with subtitles.

Table 4.2 Example of multimodal transcription with subtitles

Frame No	Time (h:m:s)	Frame	Visual Image	Kinesic Action	Soundtrack	Subtitles (BT)
1	0:11:36		CP: stationary CD: CS SA VC: elegant clothes ED: continuous VF: Sherlock keeps direct eye contact with Watson	(Sherlock is standing straight. He is talking to Mary but turning his head in the direction of Watson)	SH: I do not know why it took him so long to introduce us properly. P, m, MP	SH: I cannot understand why he so procrastinated to introduce us.
2	0:11:37		CP: stationary CD: CS SA VC: Sherlock's elegant clothes ED: continuous VF: Sherlock keeps direct	(Sherlock is standing straight. He is talking to Mary but turning his head in the direction of Watson)	SH: I do not know why it took him so long to introduce us properly. P, m, MP	SH: I cannot understand why he so procrastinated to introduce us.
3	0:11:38		CP: stationary CD: CS SA VC: Sherlock's elegant clothes ED: continuous VF: Sherlock keeps direct eye contact with Watson	(Sherlock is standing straight. He is talking to Mary but turning his head in the direction of Watson)	SH: I do not know why it took him so long to introduce us properly. P, m, MP	SH: I cannot understand why he so procrastinated to introduce us.
4	0:11:39		CP: stationary CD: CS SA VC: elegant clothes ED: continuous VF: Sherlock keeps direct eye contact with Watson	(Sherlock is standing straight. He is talking to Mary but turning his head in the direction of Watson)	SH: I do not know why it took him so long to introduce us properly. P, m, MP	SH: I cannot understand why he so procrastinated to introduce us.

In SH1_SUB (Table 4.2), Sherlock is mocking Dr. Watson through an ironic comment about the latter's future wife. Table 4.2 gives a detailed overview of different semiotic resources that concurrently construe multimodal irony. For instance, a close up shot and a stationary position of the camera let the viewers see Sherlock's face clearly. In this way they can notice some characteristic features of Sherlock's facial expressions such as wide opened eyes, raised eyebrows and "a blank face" in frame 1 and 2 or Sherlock's smirking and wry smiling in frame 3 and 4. In addition, viewers should perceive Sherlock's flat intonation and

instrumental music in the background as subtitles give viewers full access the on-screen character's original tone of voice and music, which may enhance the retrieval of ironic meaning in this excerpt. The next example in Table 4.3 illustrates the construal of multimodal irony and relay with voice-over.

Table 4.3 Example of multimodal transcription with voice-over
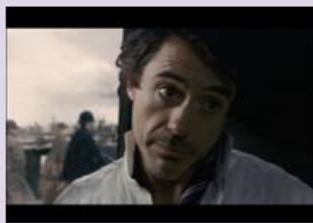
Frame No	Time (hh:mm:ss)	Frame	Visual Image	Kinesic Action	Soundtrack	Voice-over (BT)
1	0:11:36		CP: stationary CD: CS SA VC: - ED: continuous VF: Sherlock keeps a direct eye contact with Watson	◇ Sherlock pulls back his head towards the right side looking on Dr Watson from above.	[☼silence] SH: You have the grand gift of silence, Watson p, m, mp, ←→	SH: Your grand gift of silence
2	0:11:37		CP: stationary CD: CS SA VC: - ED: continuous VF: Sherlock keeps a direct eye contact with Watson	◎ △ ◇ Sherlock raises his eyebrows and keeps his eyes wide open.	[☼silence] SH: It makes you quite invaluable p, m, mp, ←→	SH: make you an invaluable
3	0:11:38		CP: stationary CD: CS SA VC: - ED: continuous VF: Sherlock keeps a direct eye contact with Watson	◎ △ ◇ (Sherlock is turning his head on left is looking at Watson from below.)	[☼silence] SH: as a companion. p, m, mp, ←→	SH: companion

Table 4.3 presents the instance of multimodal irony (SH1_VO) described more generally in Example 3.1. In contrast to SH1_SUB, here the spoken dialogue has been transferred via the voice-over, and thereby viewers can pay their full attention to the centre of the screen. In addition, a close shot and a stationary position of the camera in the combination with continuity editing allows audiences to focus on Sherlock's facial expressions (e.g., wide-open eyes, raised eyebrows, "a blank face") as well as head movements. The perception of Sherlock's tone of voice and flat intonation is only possible to some extent as the voice-over narrator covers the vast majority of the original dialogue.

As shown in Tables 4.2 and 4.3 above, the adapted model of multimodal transcription is very useful and effective as it provides a transparent and detailed overview of the verbal and non-verbal resources of film language and shows how they work in unison to produce ironic meaning on screen. The multimodal apparatus also clearly outlines which resources

play a more dominant or supportive role in this process, e.g., whether multimodal irony is construed mainly visually or acoustically. In addition, we also learn how irony is relayed with subtitles and voice-over, which is particularly important in the light of this study. Having this in mind, the multimodal transcription tool helped me identify the subtitled and voiced-over clips featuring the presence of multimodal irony that have been selected to conduct the analytical analysis in this project, as the next section illustrates.

4.5.2 Procedure for the selection of clips featuring multimodal irony

Drawing on the adapted model of multimodal transcription, the excerpts featuring the use of multimodal irony have been singled out from SH1 and SH2 to be used in the current study. A threefold selection procedure has been implemented into the decision-making process.

To start with, ten excerpts featuring the use of multimodal irony were identified by the researcher and then presented to a panel of five individuals to obtain preliminary feedback. The panel consisted of five native English speakers who were asked to watch ten pre-selected videos. The panel members pointed out these clips in which they retrieved the intended ironic meaning and recognised the importance of verbal and non-verbal elements in the construal of multimodal irony. The feedback given by the panel members resulted in reducing the number of clips from ten to eight. Since the time scheduled for a single eye-tracking recording should not exceed 30 min per participant (Webb and Renshaw, 2008), two more clips had to be excluded from the present experiment. Thus, these eight preselected videos were analysed with the adapted model of multimodal transcription. This procedure allowed me to identify how non-verbal resources are intertwined in the multimodal construal of irony on screen. As a result of this, two clips were discarded from the data set due to a smaller number of non-verbal resources construing irony. As a result, six excerpts out of eight were included in the current experiment. Special care was also taken to select videos with self-contained scenes so that the participants would not have to rely on familiarity with previous sequences of the film to understand the excerpt under analysis. The choice of the clips was further supported by the use of the two modalities of film translation, namely, subtitles and voice-over. That is to say, the multimodal transcription analysis revealed the excerpts in which multimodal irony was predominantly construed visually or acoustically. Both subtitles and voice-over have a great impact on the reception of meaning conveyed with a range of non-verbal resources, and thus can either facilitate or interfere with the retrieval of multimodal irony on screen. For instance, viewers could miss some information relayed through the visual and kinesic modes when translated with subtitles. They might also

not have the full access to the on-screen character's voice or film score due to the presence of the voice-over narrator hampering the original soundtrack. As a result, 3 clips were selected in which multimodal irony was predominantly relayed via the acoustic modes and other 3 clips in which multimodal irony was mainly relayed with the visual and kinesic modes, as shown in Table 4.4 below, in order to see whether and/or to what extent the participants are able to spot the non-verbal resources construing the ironic meaning in the presence of subtitles and voice-over.

Table 4.4. Selection of the clips in which multimodal irony was construed either through the visual, kinesic or acoustic modes in the subtitled and voiced-over clips

Clip 1 (SH1_1)	Clip 2 (SH1_2)	Clip 3 (SH1_3)	Clip 4 (SH2_1)	Clip 5 (SH2_2)	Clip 6 (SH2_3)
Acoustic modes	**Visual and kinesic modes**	**Acoustic modes**	**Visual and kinesic modes**	**Visual and kinesic modes**	**Acoustic modes**
While subtitles can support, voice-over can interfere with non-verbal resources conveyed through the original soundtrack.	While voice-over can support, Subtitles can interfere with non-verbal resources conveyed through mise-en-scène, cinematography or editing.	While subtitles can support, voice-over can interfere with non-verbal resources conveyed through the original soundtrack.	While voice-over can support, Subtitles can interfere with non-verbal resources conveyed through mise-en-scène, cinematography or editing.	While voice-over can support, subtitles can interfere with non-verbal resources conveyed through mise-en-scène, cinematography or editing.	While subtitles can support, voice-over can interfere with non-verbal resources conveyed through the original soundtrack.

As it can be seen from Table 4.4, in the clips 1, 3 and 6 the acoustic modes including music, quality of voice and other non-verbal sounds can be hardly accessible in the voice-over version due to the narrator's voice reducing the audibility of the original soundtrack, while these can be easily accessible with subtitles. On the other hand, in the clips 2, 4 and 5 multimodal irony has been principally construed through a combination of the visual and kinesic modes. The presence of subtitles can here interfere with the retrieval of multimodal irony when viewers have to divide their attention between the centre and the bottom of the screen. Meanwhile, they can fully concentrate on these elements in the voiced-over version.

In the context of this thesis, two different types of audiovisual translation modalities were added to the selected sequences, namely, subtitles and voice-over. The professional versions of the subtitled and voiced-over variants were extracted from the DVD of SH1 and SH2 according to the following procedure: The DVD was converted into VOB files with DVD Shrink (a freeware tool) to back up a DVD discs and choose what soundtrack and subtitles should be exported. Thereby, two versions of SH1 and SH2 were exported, i.e., one with subtitles and the other with voice-over in Polish. Both recorded versions were then uploaded to Freemake Video Converter (a freeware conversion tool) to cut the selected clips

from the whole source material. The clips were saved in AVI file as required by the eye-tracking software SMI Experiment Center. The same clip was displayed with both subtitles and voice-over to two parallel groups of participants, and therefore two main groups, namely SUB_VO and VO_SUB, were created, as shown in Table 4.5.

Table 4.5. Experimental design with two groups of respondents watching subtitled and voiced-over clips

Type of Translation	SUB_VO	VO_SUB
Subtitles/voice-over	3 subtitled clips	3 voiced-over clips
Voice-over/subtitles	3 voiced-over clips	3 subtitled clips

As illustrated in Table 4.5 subjects from SUB_VO group first watched 3 clips with subtitles and then 3 clips with voice-over, while subjects from VO_SUB groups started watching the same clips but in the reversed type of translation. This experimental design allowed me to bring out the exact differences between the two groups. That is to say, it was possible to compare how the subjects comprehended the same multimodal content with subtitles and voice-over and whether and/or to what extent they were able to retrieve these verbal and non-verbal modes of film language while watching the clips under scrutiny featuring the use of multimodal irony.

4.6 Interactionist Component: Questionnaire

This section examines questionnaire as a methodological tool to elicit responses from the respondents in this experiment. Some general rules are explored that should be considered when designing a questionnaire particularly in the light of the present study. A completed questionnaire from the data set has also been presented to illustrate the proposed model of coding procedure to evaluate the participants' responses.

4.6.1 Questionnaire as a method of data collection

Questionnaires to date have been widely used in the reception studies of audiovisual content (e.g., Jensema, 1998; Fuentes Luque, 2003; Bucaria and Chiaro, 2007; Caffrey, 2009; Desilla, 2009; Schauffler, 2012; Ramos Pinto, 2013), for instance, in the perception of non-verbal items in audiovisual material (Bucaria, 2005) or in the comprehension of implicature in subtitled clips (Desilla, 2009). More importantly, the combination of questionnaire and eye-tracking data is becoming more and more common among AVT researchers to examine the perception of pop-up glosses in TV anime (Caffrey, 2009), to collect data on the

participants' cognitive effort and performance while watching standard subtitled clips and standard subtitled clips with surtitles (Künzli and Ehrensberger-Dow, 2011) or to test the understanding and perception of professional and non-professional subtitles (Orrego-Carmona, 2015), among others. The usefulness of this research method in the aforementioned studies supports its inclusion in the current experiment to complement the eye-tracking data and to probe the comprehension of the intended ironic meaning in the selected clips of SH1 and SH2. A questionnaire has also been used to determine what elements might have helped the participants reach the conclusion that the meaning is intended to be ironic.

Apart from the questionnaire testing the comprehension of multimodal irony in the present study, two additional questionnaires have been included in the current project, that is, the pre- and post-experimental questionnaire (see Appendix 2 and 4). In the pre-experimental questionnaire, the respondents were asked to complete a self-assessment questionnaire which was designed to assess their level of proficiency in English in order to allocate them to the relevant language group. The subjects' level of proficiency of the second language can be measured in at least two ways before participants are assigned to a certain language level subgroup. A listening-comprehension test appears to be one of the most reliable methods to ensure that some internal linguistic consistency among participants is maintained (e.g., Orrego-Carmona, 2015). Nevertheless, the use of self-assessment questionnaires can bring out equally reliable results, if the questions inquiring the five key factors such as formal tuition, level of immersion, active use, passive use and confidence are considered in the design (Schauffler, 2012). According to Schauffler (2012), each question considered separately is insufficient to make legitimate claims about the participants' level of proficiency in the second language but "it is the combination of the five which is hoped to give a near adequate impression as to the proficiency of the person" (Schauffler, 2012: 138). As a result, the successful use of self-assessment questionnaires including the above-mentioned key factors in the previous studies supports its inclusion in the present research of multimodal irony to test the level of the source language proficiency of the viewers. The pre-experimental questionnaire also verified the participants' knowledge about Sherlock Holmes as a Sir Arthur Conan Doyle's fictional icon and the two films under study. In the post-experimental questionnaire, the subjects were asked to answer a multiple-choice question. In this question, they were provided with a list of non-verbal elements of the language of film and were asked to select these elements that, in their opinion, contributed to their comprehension of multimodal irony. Subsequently, their demographic data (age and gender)

were collected and information about the university course they studied. The participants were also asked about their audiovisual consumption habits and audiovisual translation preferences when watching the audiovisual content to identify a profile of the Polish viewer (Gambier, 2003, Orrego-Carmona, 2015).

In the next section, several important aspects are discussed that need to be considered when designing a questionnaire probing the comprehension of irony in the selected clips under analysis.

4.6.2 Questionnaire design for the present study

Questionnaire design depends on the research objectives and the type of information the researcher wishes to collect, namely qualitative and quantitative, to answer the research questions. Without a doubt, the wording, length, structure of a questionnaire as well as the type of questions (e.g., open-ended, closed or multiple choice) have to been taken into account when writing a questionnaire. There are several facets that have been considered important while writing the comprehension questionnaire in the context of the present study.

Wording. In order to increase the completion rate and maximise the precision of responses, jargon or specialised terminology was excluded to eliminate the feeling of ambiguity and confusion in participants (e.g., O'Brien and Saldanha, 2014; Wilson, 2013). Thus, more complex terms like *non-verbal semiotic resources* were replaced with more common phrases such as *elements* so that each question conveys the same meaning for everyone (Adams and Cox, 2008). To decrease the non-completion rate or the number of skipped words, several questions were rephrased to maintain the participants' attention and prevent automatic responses (O'Brien and Saldanha, 2014).

Length. Both the length of questions and the whole questionnaire should be considered. Lengthy questions were avoided to eliminate misinterpretations or inaccurate responses (Adams and Cox, 2009). The comprehension questionnaire was also kept as short as possible due to respondents' general short attention span (O'Brien and Saldanha, 2014). As a result, the subjects in the present study were asked to answer four questions after each clip probing their irony comprehension in order not to overload them with abundant information in one go (O'Brien and Saldanha, 2014).

Type of questions. It is also vital to decide the types of questions that are to be employed in the questionnaire, that is, open-ended, closed or multiple-choice questions. To test the comprehension featuring the intended ironic meaning, open-ended questions were employed to extract fuller and richer information from participants, to explain the reasons

behind their answers or to elaborate further on their choice (e.g., O'Brien and Saldanha, 2014). The open-ended questions thus allow respondents to "add further opinions or to highlight an opinion they hold and which is not addressed in the questionnaire" (O'Brien and Saldanha, 2014: 157). Additionally, a combination of closed or multiple-choice questions was incorporated to collect precise answers and complement the participants' responses given to the open-ended questions.

Structure. The questionnaire should be logically structured so that it sounds "like a story" starting with simple questions and followed by more complex ones (Rasinger, 2008: 70-71; O'Brien and Saldanha, 2014). In the present study, the comprehension questionnaire started with introductory scenarios to each clip to help the participants contextualise their responses. Once they watched the video clip, their irony comprehension was examined in the first two questions. In the third question, it was explored what elements helped them come to the conclusion that the intended meaning is believed to be ironic. The last question aimed to gain insight into the retrieval of multimodal irony in the subtitled and voiced-over clips under scrutiny.

Reliability and validity. Last but not least, the concepts of reliability and validity are essential to the questionnaire design. Reliability here "refers to the consistency of a measure whilst validity refers to its ability to measure what it is supposed to be measuring" (Adams and Cox, 2009: 18). In the present study, reliability refers to the congruity on the responses about irony comprehension given by many participants over and over again, while validity refers to the degree the question actually measures irony comprehension.

4.6.3 Questionnaire analysis and coding procedure in the present study

As mentioned in the subsection above, the questionnaire probing irony comprehension in this study included four questions. Prior to the video display, the participants read a brief contextual introduction to the scene. Straight after the video display, they were asked four questions (Q1; Q2; Q3; Q4) in relation to the clips they saw. Due to the type of questions selected for the questionnaire, namely, open-ended, the participants were asked to give their responses in the box below each question. The questionnaire has been outlined in Table 4.6 below.

Table 4.6 Questionnaire examining irony comprehension in the subtitled and voiced-over clips of SH1 and SH2 used in the experimental part of the present study

Question	Polish version	Back translation in English
Q1	Dlaczego Watson zwraca się do Sherlocka mówiąc "Wyglądasz cudownie"	Why is Watson saying to Sherlock: "You look gorgeous"
Q2	Jaki byś określił(a) sposób w jaki Watson zwraca się do Sherlocka w tej scenie?	How would you describe the way in which Watson is speaking to Sherlock in this scene?
Q3	Czy zauważyłeś (-łaś) lub usłyszałeś (-łaś) podczas tej sceny elementy takie jak gesty, ruchy ciała, scenerię które ułatwiły Ci zrozumieć to co Watson miał na myśli?	Have you noticed or heard during this scene any gestures, body movements, scenery that helped you understand what Watson was trying to say?)
Q4	Czy myślisz, że tłumaczenie z lektorem/napisy przeszkodziło Ci w jakiś sposób zrozumieć to co Sherlock chciał wyrazić w tym fragmencie? (Jeśli wybrałeś odpowiedź 1 lub 2 proszę uzasadnij w polu poniżej)	Do you believe that the translation with voice-over/subtitles hindered somehow your understanding of what Sherlock wanted to say in this fragment? (If you chose the answer 1 or 2 please justify in the box below)

As presented in Table 4.6, first, I gauged the viewers' irony comprehension in Question 1 (Q1) and Question 2 (Q2). Hence, the participants were first asked (Q1) to explain why Watson says the following to Sherlock in the given scene. Q1 also included the relevant dialogue to indicate the exact moment in the clip under analysis the participants should focus on to give more precise and meaningful responses. To enhance the validity of the questionnaire, the second question (Q2) additionally re-checked the retrieval of multimodal irony asking them to assess Watson's attitude towards Sherlock from their own perspective in regard to the particular moment indicated in Q1. Following Q1 and Q2, the respondents were asked the third question (Q3) that examined how they have come to believe that the presented meaning was intended as ironic. Explicitly, I was interested to see what non-verbal resources of the language of film such as gestures, facial expression, music score or camera perspective, to name just a few, contributed to their understanding of multimodal irony. The last question (Q4) was tailored to assess whether and/or to what extent subtitles or voice-over interfered with the participants' comprehension of irony in the selected clips of SH1 and SH2. Q4 was supposed to indicate whether and/or to what extent voice-over limited their access to sounds. In Q4 my intention was to see whether and subtitles dragged their attention away from the non-verbal resources visible on the screen. Questions assessing translation difficulty have already been successfully implemented in other eye tracking studies exploring the comprehension of the subtitled clips (e.g., Orrego-Carmona, 2014b).

In order to interpret the responses and measure whether/to what extent the subjects understood the intended meaning in Q1 and Q2, a coding procedure has been designed drawing on Strauss and Corbin's (1990) grounded theory defined by Adams et al. (2009: 139) as "a method of qualitative research that aims to produce new theories […] grounded in the qualitative data gathered during the research". Strauss and Corbin (1990) suggested that grounded theory is particularly useful when analysing complex phenomenon like irony. They put forward three different stages of the coding process, i.e., selective, open and axial, which often intersect to provide a researcher with the most suitable coding procedure. In the light of this study, the coding procedure was incorporated in the following way. In the first stage, the central phenomenon under study was identified (i.e., core category), which is multimodal irony in the current study. In the second stage, categories revealing the characteristic properties of the category were recognised. In respect to my project, attitude of dissociation was classified as a category with its characteristic properties such as "speaker's mocking, scornful or contemptuous attitude" (Wilson, 2013: 53). In the third stage, the participants were asked to explain the conditions that lead to the retrieval of the phenomenon under scrutiny, i.e., multimodal irony. This coding procedure was linked with a purpose-built scale similar to those devised to assess students' reading comprehension skills, which has been successfully applied in the previous studies (e.g., Desilla, 2009). The scale in Table 4.7 presents the number of points that the participants received in regard to the quality and relevance of their responses to Q1 and Q2.

Table 4.7 A scoring scale corresponding to the level of relevant information given in the questionnaire responses

Score	Description
0	• **Either** providing a completely inaccurate or irrelevant answer **or** no answer at all.
1	• Obscure, inconclusive evidence of accessing the intended ironic meaning.
2	• Understanding irony as a core category explaining its nonliteral meaning • **Either** understanding its relative concepts **or** providing a description of attributes related to ironic meaning.
3	• Understanding irony as a core category explaining its nonliteral meaning • **Either** understanding its relative concepts **or** providing a description of attributes related to ironic meaning. • Explaining the conditions that lead to the construal of the ironic meaning

As shown in Table 4.7, the participants were awarded no points when they provided no or blatantly incorrect and inconclusive responses; one point was awarded for responses indicating superficial understanding of ironic meaning, while additional 2 points were given to the responses reflecting the retrieval of the core category accurately along with its relative concepts or attributes. The additional 3 points were granted when the respondents explained the circumstances that contributed to the construal of ironic meaning in addition to the identification of the core category and its related concepts. As a result, each participant could get maximum 6 points in total.

The coding procedure in regard to Q3 and Q4 was much simplified. That is, in response to Q3 the number of participants was counted in each subgroup that mentioned at least one of the non-verbal elements pertaining to the visual, acoustic and kinesic modes of the language of film. Responses to Question 4 were coded using a multiple-choice format: 1-No; 2 - Rather no; 3 - I don't know; 4 - Rather yes; 5-Yes. As a way of illustration of the coding apparatus, the questionnaire has been presented in Table 4.8 with the corresponding responses collected during the experiment. Each participant was assigned an individual code consisting of a capital letter P and a number, while P stands for a participant, numbers 01, for instance, specifies a participant's number in order to anonymise the collected data. I assigned a code to each individual prior to the commencement of the eye-tracking study.

Table 4.8 Questionnaire with three questions probing irony comprehension (1, 2) and non-verbal elements construing meaning (3) with two examples of participants' responses from the experimental study

No	Question	Response in Polish	Back Translation
Q1	Dlaczego Watson zwraca się do Sherlocka mówiąc "Wyglądasz cudownie" ? (BT: Why is Watson saying to Sherlock: "You look gorgeous"?)	P18A: "Myślę, że mówi to ironicznie, widzi jak Sherlock budzi się z potarganymi włosami i wcale nie myśli, że wygląda cudownie."	BT: "I think he is saying this ironically, he sees Sherlock waking up with tousled hair and he doesn't really think that he looks gorgeous."
		P21B: "Używa sarkazmu (moim zdaniem), pokazuje to również w pewnym sensie, że są oni dobrymi znajomymi."	BT: "He is using sarcasm (in my opinion) and shows in a way that they are good friends."
Q2	Jaki byś określił(a) sposób w jaki Watson zwraca się do Sherlocka w tej scenie? (How would you describe the way in which Watson is speaking to Sherlock in this scene?)	P18A: "Ironiczny"	BT: "Ironic"

		P21B: "Wydaje się, że Watson bardzo dobrze zna zachowania oraz nawyki Sherlocka. Nie dziwi go jego zachowanie."	BT: "I guess that Watson knows Sherlock's habits very well. He is not surprised with his behaviour."
Q3	Czy zauważyłeś (-łaś) lub usłyszałeś (-łaś) podczas tej sceny elementy takie jak gesty, ruchy ciała, scenerię które ułatwiły Ci zrozumieć to co Watson miał na myśli? (BT: Have you noticed or heard during this scene any gestures, body movements, scenery that helped you understand what Watson was trying to say?)	P18A: "Wydaje mi się, że miał drwiący usmiech gdy to mówił. Poza tym jego wyprostowana postawa."	BT: "I think he had a mocking smile when he was saying this. Apart from this his straight posture."

The responses given to questions 1-3 by P18A and P21B in Table 4.8 illustrate how participants' responses to question Q1, Q2 and Q3 were qualified. In regard to Q1, both P18A and P21B responded correctly yet P18A was scored 3 points and P21B 2 points according to the coding procedure presented in Table 4.7. In this case, both participants encoded ironic meaning correctly yet P18A additionally explained the circumstances that lead to the retrieval of the ironic meaning. In Question Q3, these participants were identified who mentioned at least one non-verbal element that helped them come to the conclusion that the meaning may be considered ironic.

In the current research project, a questionnaire has been incorporated in the methodological design as complementary data source to eye-tracking data to "gain insight into participants' experienced cognitive processes while inspecting a stimulus or performing a task" (Holmqvist et al., 2011: 96). In the next section, I therefore present the eye tracking technology as a methodological tool to enhance my understanding of where, when and how frequently the viewers looked at to retrieve meaning, in particular ironic meaning, in the selected clips of SH1 and SH2.

4.7 Experimental Component: Eye Tracking

This section gives an account of eye-tracking technology and discusses how eye-tracking research can be brought into dialogue with film and audiovisual translation studies in the context of the present study. In the last subsection, it is explained how this technology was used to conduct the experimental part of this research project.

4.7.1. Understanding eye-tracking technology

Currently, eye tracking represents a true technological revolution. The eye tracker consists of a high-resolution camera that utilises the near-infrared technology to track the direction of the eye drawing on the concept referred as Pupil Centre Corneal Reflection (PCCR). It means that near-infrared light is flashed into the pupil causing visible reflections in the cornea which are mirrored back and tracked by the camera. Eye tracking thus records and measures the eye activity in real-time on 2D screens or in 3D and real-life environments in a non-obtrusive way providing researchers with objective data as to the position and movement of the eye on a stimulus.

The eye-tracking technology is therefore particularly useful to find out how viewers "read complex text, what attracts their attention and what does not and how they integrate information from the language, images, sounds, animations" (Holsanova, 2014: 288). In the context of the current study, the application of eye tracking technology is also very advantageous as it enables me to bring out differences in gaze patterns and attention dispersal of my participants when watching "very complex dynamic environments as encapsulated in movies" with subtitles and voice-over featuring multimodal irony (Dyer and Pink, 2014). In addition, eye tracking also allows me to examine the extent to which a level of language skills affects the way the participants retrieve meaning in the selected clips under scrutiny especially when dialogues are not literally communicated and "the viewer must then use other cues to work out the importance of the speech" (Robinson et al., 2015).

Regarding the types of eye trackers available, there are remote, mobile and head-mounted systems. Remote eye trackers mounted to the desktop monitor are the most commonly used in the study of any screen-based stimulus material (e.g., pictures, videos or websites). Because of the nature of the material used in the present study, that is, video-based content, a remote eye tracker was employed (SMI RED250) at 120Hz by SensoMotoric Instruments GmbH (SMI) to conduct the experimental part of this project, as presented in the Figure 4.3. A sampling rate of 120 Hz appears to have become a standard frequency used in AVT eye tracking studies (e.g., Kruger 2012; Orrego-Carmona, 2015).

Figure 4.3 SMI RED250 at 120Hz used in the present study

Eye Tracker

SMI RED250 (Figure 4.3) does not require participants to keep their head completely still with the help of a bite bar or a chin rest and allows for relatively free head movements within 32 x 21 cm at 60 cm distance. This enables to record the eye movements in a more naturalistic environment.

There are two main types of eye movements that enable the interpretation of the data: **"fixations"** and **"saccades"** (O'Brien, 2010). A fixation refers to the position of the eye in which it remains relatively still over a certain period of time (approx. 200-300 ms), for instance, when participants keep looking at a word while reading (Holmqvist et al., 2011). Fixations are thus used by the majority of researchers to analyse and understand viewers' visual and cognitive processes in different types of stimuli across a wide array of disciplines (Webb and Renshaw, 2008; Redmond and Batty, 2014; Robinson et al., 2015; Holsanova, 2008) including the research on the use of subtitles (e.g., Kruger, 2012; Szarkowska et al., 2014; Orrego-Carmona, 2015). The ballistic eye movement, on the other hand, that rapidly drifts from one fixation point to another (approx.30-80 ms) is termed as a saccade. Saccades are also widely used and reported in the research, for example, in studies involving reading activity (Holmqvist et al., 2011).

In order to obtain accurate interpretations of eye movements, researchers use areas of interest (AOIs) as a tool for further quantitative analysis of the eye-tracking data. The AOIs can be explained as "researcher-defined areas within the visual stimulus" (Hall et al., 2013: 38) used to quantify when and how many times respondents look at the particular region of stimulus. There are two types of AOIs, namely, static and dynamic, that are applied to different kinds of stimuli. While static AOIs are fixed with well-defined borders drawn on an image, dynamic AIOs change their shape, size and position over time and in reference

to the screen (Papenmeier & Huff, 2010). Dynamic AIOs thus represent an ideal solution for the analysis of semantically important content in the multimodal texts like animated stimuli or films with subtitles (e.g., Jensema et al., 2000; Duchowski, 2007; Kruger, 2012; Szarkowska et al., 2014), to name just a few. For this reason, dynamic AOIs have been employed in the present study of multimodal irony and drawn on the on-screen characters' faces, bodies and in the subtitle area in SMI BeGaze software, as presented in Figure 4.4.

Figure 4.4 AOIs drawn around faces and bodies on the on-screen characters as well as around the subtitles in SMI BeGaze software in the SUB_VO and VO_SUB groups

SH1_1   Voice-over   Subtitles

SH1_2

SH1_3

103

SH2_1

SH2_2

SH2_3

As presented in Figure 4.4, there is a separate AOI delineated accurately around the on-screen characters' faces and body in the voiced-over excerpts. I did not use typical geometrical figures (e.g., a rectangle or circle shape) to draw an AOI around face and body but used the connecting points to delineate an AOI around these areas of interest as precise as possible. In addition, a separate AOI was drawn around the subtitle lines when these appeared on screen. In the case of two-line subtitles, two separate AOIs were drawn around the first and second subtitle line. In the analysis of the eye-tracking data, the fixation values from two separate AOIs around two-line subtitles were added up and presented as one value. The eye-tracking data was recorded in the given AOIs only in the presence of the ironic dialogue, that is, out of the whole clip duration there were only a couple of seconds when the eye movements were recorded. The dynamic AOIs adjusted their shape and size automatically in reference to the on-screen character's movements. The duration of the

104

whole recording and the duration of the recorded eye movements in the AOIs have been presented in Table 4.9 below.

Table 4.9 The duration of videos and the recorded eye-tracking (ET) data in the experiment

Clips	Duration of video	Duration of ET data
SH1_1	00:01:15	10 s
SH1_2	00:01:17	5 s
SH1_3	00:02:35	8 s
SH2_1	00:01:29	14 s
SH2_2	00:01:23	12 s
SH2_3	00:05:07	20 s

In the context of this study, the number of fixations, total fixation duration as well as mean fixation duration per AOI have been used to interpret the eye-tracking data which appear to be a standard in the research of multimodal and audiovisual texts (Holmqvist et al. 2011; Kruger, 2012; Kurger and Steyn, 2014; Orrego-Carmona, 2015).

The total fixation duration is calculated in milliseconds by adding all the fixations that were recorded in a given AOI and reflects the time that all participants (HLPs, MLPs, LLPs) spent in this AOI in total while watching the excerpt. For example, the sum of all fixation durations in the AOI around Sherlock's face in Clip SH1_1 for HLPs in the SUB_VO group resulted in the total duration of fixations in this AOI of 5610 milliseconds (ms). The total fixation duration is also presented as the percent ratio which was calculated between the total fixation duration in a given AOI and the total gaze time recorded in the AOIs in the given clips. For instance, if the total fixation duration in the AOI around Sherlock's face per group of participants in Clip SH1_1 (e.g., HLPs in the SUB_VO group) is 5610 ms and the total fixation duration across all AOIs in this video for the same HLPs subgroup is 15556 ms which equals to 100%, then HLPs allocated 36% of their attention to the AOIs Sherlock's face.

The number of fixations was also determined as the percent ratio which represents the value between the total number of fixations in a given AOIs and the total number of fixations in the AOIs in the recording. That is to say, if HLPs in the SUB_VO group had 15 fixations in the AOI around Sherlock's face and 65 fixations in the AOIs in Clip SH1_1 which equals 100%, then HLPs would have fixated 23% on Sherlock's face in the recording.

The eye-tracking data was additionally analysed in terms of the mean fixation duration calculated for each subgroup (HLPs, MLPs, LLPs) in the SUB_VO and VO_SUB group in each clip to reveal the average duration of fixations (in ms). The mean fixation duration was computed by dividing the sum of the duration of fixations by the number of

fixations per subgroup. Since the mean fixation duration is considered a central statistic parameter in reception studies of the film texts examining the relationship between cognitive processing and visual behaviour, I used these fixations values to conduct statistical tests.

Estimating the total fixation duration and the number of fixations as the percent ratio in tandem with the calculation of the mean fixation duration, the areas of interest could be identified that the participants looked at the most frequently, and thus compare the ways they behaved when watching the voiced-over and subtitled clips. That is to say, longer fixations on subtitles may reveal higher cognitive efforts as the fixated AOI may require from the participants, for example, with low level of English proficiency, more time to process the information enclosed in the AOI. The participants with high level of English proficiency, on the other hand, may look at subtitles rather incidentally, and use them to reconfirm the information they missed or vocabulary they misunderstood in the original soundtrack (Orrego-Carmona, 2015). Although the fixation-based metrics have been principally selected as a source of cognitive and attentional information for the current project, there is also one more eye-tracking data used in the field of audiovisual translation (e.g., Caffrey, 2009), namely, pupillometrics. Although this metric might bring new insights to the analysis, it would be very difficult to control it in this project due to the pupils' sensitivity to light flashing from videos with different colours and with varied intensity. In addition, a wide range of other stimulants such as stress, emotions, caffeine or alcohol might also affect the pupil size. As a result, it would be challenging to define what exactly causes the change in the pupil size (Holmqvist et al., 2011; Hvelplund, 2014).

To visualise eye-tracking data and complement the quantitative analyses, it is also possible to include visualisations such heat maps or scan paths. A heat map (Figure 4.5) represents "spatial distribution of data" which, in other words, simply shows where participants look (Holmqvist et al., 2011: 238). A scan path (Figure 4.6) represents the paths of the eyes movement when scanning, viewing and analysing visual information in the visual field (Holmqvist et al., 2011).

Figure 4.5 Heat map from SH2_1   **Figure 4.6** Scan path from SH1_1

Despite their apparent advantages, heat maps and scan paths are excellent visualisations solely for the first quality check of eye-tracking data to see whether the fixation detection algorithm performed well (Holmqvist et al., 2011). One should stay cautious when interpreting heat maps and scan paths data as they "give an easily digestible overview of the total data from a large number of participants" (Holmqvist et al., 2011: 239) and their interpretation is very misleading as they hide a whole range of details and "do not lend themselves to any systematic comparison" (Bojko, 2009: 37). In the present study, these visualisations have been therefore only used to illustrate viewing behaviour of the participants at different levels of English when watching the subtitled and voiced-over clips.

Eye-tracking technology is used in a wide range of disciplines from neuroscience, psychology and human factors to linguistics (Holmqvist et al., 2011). More recently, the eye-tracking technology is also utilised in the reception of multimodal texts like films (e.g., Smith, 2014; Redmond et al., 2015) and the subtitled films (e.g., Kruger, 2012; Szarkowska et al., 2016), as the following section reports.

4.7.2 Eye tracking in the reception of on-screen multimodal texts

Eye-tracking studies are grounded on the eye-mind hypothesis put forward by Just and Carpenter (1976). Their hypothesis states that "people look at what they are thinking about" (Webb & Renshaw, 2008: 39) and their visual behaviour is influenced by their reactions to bottom-up stimuli (e.g., triggered by visual information like flickering colours) and top-down processes (e.g., triggered when a person activates cognitive efforts to comprehend textual information), when the brain is engaged in the processing of the visual field.

Although it is necessary to combine eye tracking with other methodological tools (like questionnaires) to distinguish between what people see vs. perceive, the general principle of the eye-mind hypothesis can be applied to AVT research in which viewers have to split attention between reading the subtitles and perceiving the information in the centre part of the screen. That is to say, the viewer's fixations on subtitles may indicate the cognitive processing of verbal resources, while fixations on the reaming parts of the screen may mean the processing of visual resources.

Eye tracking has thus become a well-established tool to measure both attention allocation and cognitive load in a wide array of multimodal texts (Kruger and Steyn, 2013). For instance, eye tracking studies have significantly flourished in the domain of film studies in recent years "as a key to successful filmmaking" (Brown, 2014) supporting film directors and visual artists to redefine their craft (Treuting, 2006) and enhancing visual experiences

in the spectators (Redmond and Batty, 2014). Hence, eye tracking is very effective to measure how audiences comprehend meaning relayed through the visual, kinesic and acoustic modes. For example, in regard to the visual modes, gaze patterns of viewers are affected by certain camera movements and perspectives (Redmond et al., 2014) and by other visual resources pertaining to mise-en-scène such as objects, clothes or furniture (Marchant et al., 2009). Editing patterns are also seen to affect eye movements during scene perception (Smith and Mital, 2013; Redmond et al., 2015). For instance, Smith (2012) found out that continuity editing[23] is the most common strategy to grab viewers' attention. The kinesic modes have also been found as important especially when viewers dwell "back and forth between the eyes, face and mouth of the central characters" (Redmond et al., 2015) in the search for additional information to retrieve the intended meaning (Robinson et al., 2015). Additionally, the acoustic modes have been shown to modulate gaze patterns of viewers (Robinson et al., 2015). Coutrot et al. (2012) compared the eye-movements in the sound on and off conditions and identified several differences in fixation duration and fixation location in the clips with and without sound.

Eye tracking has also been productively employed in the field of AVT research with an attempt to investigate attention to and processing of subtitles (e.g., Specter, 2008; Szarkowska et al., 2014; Kruger, 2012; Kruger and Steyn, 2013; Perego et al., 2010; Duchowski, 2007). Since viewers have to process several sources of information simultaneously, they may experience difficulties sharing their attention between the centre and the bottom of the screen (e.g., d'Ydewalle and De Bruycker, 2007). It happens when people are not used to reading "the fleeting text on a dynamic background" with a constant speed (Kruger, 2012: 105) or rely on subtitles considerably to retrieve meaning due to a low level of proficiency in the source language (Orrego-Carmona, 2015). Among numerous eye-tracking studies, Jensema et al. (2000) studied visual behaviour of respondents watching video excerpts with and without subtitles. She indicated that viewers quickly moved their eyes to subtitles when those appeared on the screen. In another study, Kruger, Szarkowska and Krejtz (2015) revealed that viewers shift their gaze back to the centre of the screen within a few seconds after having read the subtitles. Specker (2008), on the other hand, conducted the multimodal analysis of reading the subtitles in the presence of non-verbal sources of information, while Orrego-Carmona (2015) examined how viewers with a high and low level of English process professional and non-professional subtitles.

---

[23] Continuity editing is an editing technique that to have to keep the progress of movie progress (transition) from one shot to another shot in a smooth (Cahir, 2012).

Despite a great deal of eye tracking studies on the reception of subtitles, in particular in Polish AVT research (e.g., Szarkowska et al., 2016), it is still rather unexplored how viewers process more complex meaning, in particular ironic meaning, in the subtitled and voiced-over films. Especially, in the context of Polish AVT scene, no single study exists, to the best of my knowledge, investigating irony comprehension of its audiences at three levels of English proficiency (i.e., high, medium, low) relayed with subtitles and voice-over. Despite a growing number of enthusiasts of subtitles in voice-over countries like Poland (Chaume, 2013; Szarkowska and Laskowska, 2014), voice-over is still reigns on TV in the vast majority of Baltic European countries and this trend is likely to continue due to economic and sociological reasons. Therefore, sufficient consideration should be given not only to the reception of the subtitled but also voiced-over films to determine how these viewers engage with this mode of AVT, too. In order to achieve this goal, the following eye-tracking study was designed and conducted, as described in the next section.

4.7.3 Eye-tracking pilot study and lessons learnt

The eye-tracking pilot study preceded the eye-tracking experiment to examine the proposed methodological apparatus, questionnaire and experimental design. In addition, the pilot study helped me familiarise myself with the eye-tracking equipment and gain experience in planning and running this type of empirical studies.

The pilot study took place on 17th April 2015 at Web Ergonomics Lab located at School of Computer Science (University of Manchester) and 8 Polish native speakers participated in the eye-tracking sessions on that day. All participants were graduate, postgraduate or doctorate students of the University of Manchester (UoM) with high level of proficiency in English language (the evidence of English language proficiency, such as CAE or CPE, is required to be accepted for admission to UoM). All subjects had normal or corrected-to-normal vision (by wearing glasses or contact lenses). The aim of the study was not revealed but the participants were informed that their eye movements are recorded to enhance their spectatorial experience of watching foreign movies translated with two different translation modalities, i.e., subtitling and voice-over.

6 videos were selected from *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011) and presented to the participants of the study. In the experimental design, 3 clips were displayed with subtitles and other 3 with voice-over. Subjects were asked to answer the experimental questionnaire with open-ended questions after each video was displayed. In the pre-experimental questionnaire, the demographic data was collected from

the participants. The study was designed with Tobii Studio in which AOIs were drawn in a geometrical shape around the on-screen characters' face and body and subtitle-lines. The experiment was recorded with the eye tracker Tobii X2-60 and the collected data were exported from ClearView software.

Following the pilot study, two fundamental changes were introduced to the main experiment. The most important modification refers to the experimental design. In the "mixed factorial" design of this pilot study, each participant watched three subtitled and three voiced-over clips. Since all videos were different, it seemed to be difficult to fully understand the exact discrepancies in the retrieval of multimodal irony in the subtitled and voiced-over clips. As a result, due to the different video length and varied size of AOIs in each clip, I found it hardly possible to compare the duration of fixations and the number of fixations in the subtitled and voiced-over clips, and thus interpret the eye-tracking data. I therefore decided to collect two parallel groups of mixed factorial designs in the main experimental study in which the two different groups of participants watched the same subtitled and voiced-over versions of SH1 and SH2.

The modifications related to the questionnaire design involved the addition of post-experimental questionnaire as a multiple-choice question to complement the responses given to the open-ended questions in the experimental questionnaire and to collect more accurate responses. The post-experimental questionnaire listed non-verbal resources pertaining to the language of film and the participants were asked to select as many responses as deemed necessary. The list of the non-verbal elements involved these markers of multimodal irony that were enumerated in the multimodal transcription framework in respect to the visual, acoustic and kinesic modes. The multiple-choice question enabled me to establish what non-verbal resources exactly the viewers found helpful in their comprehension process of multimodal irony.

4.7.4 Eye-tracking experiment

The eye-tracking experiment was conducted in two sessions in the AV Lab at the Institute of Applied Linguistics, University of Warsaw. The first session took place from 10th May 2015 and lasted until 28th May 2015, while the second session was scheduled from 17th September 2015 to 25th September 2015. The dates of the eye-tracking sessions were scheduled in line with the availability of the eye-tracking laboratory. A group of 56 answered to my announcements about the experiment of which 48 volunteers, who met the necessary criteria, participated in the two eye-tracking sessions. Due to calibration problems and poor

calibration values, data from 4 subjects was removed from the experiment before the eye-tracking recording was initiated, leaving a total of 44 participants. The number of participants allowed me to incorporate a quantitative dimension in my analysis (e.g., Webb and Renshaw, 2008).

Participants were recruited 6 weeks prior to the start of the first session. The information about the upcoming study was widespread through notices, adverts, social media groups, university mailing and personal networks. The adverts were spread across different campuses of the University of Warsaw, for instance, at the physics, chemistry, engineering, social sciences, economy, journalism, psychology and linguistics departments.

Prior to the recording sessions, it was ensured that the eye tracking system (iView X and Experiment Center) was set up and connected correctly. The stimulus was displayed on 23" LCD desktop monitor with the eye tracker fixed below the screen. The desktop monitor was connected to the researcher's laptop which was located on the participant's left-hand side to make my presence as a researcher as little obtrusive as possible. Figure 4.7 below illustrates the layout of the recording area, while Figure 4.8 4.9 and 4.10 shows the eye tracking laboratory in which the sessions were recorded at the University of Warsaw (UW).

Figure 4.7 Layout of the recording area   **Figure 4.8** Layout of the eye-tracking lab at UW
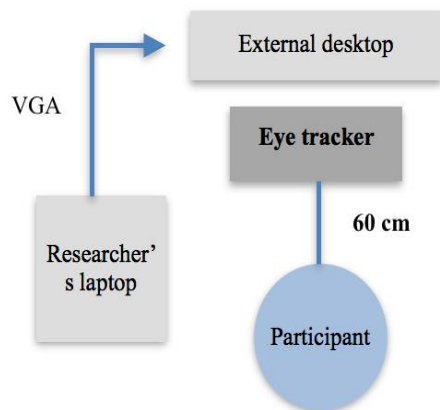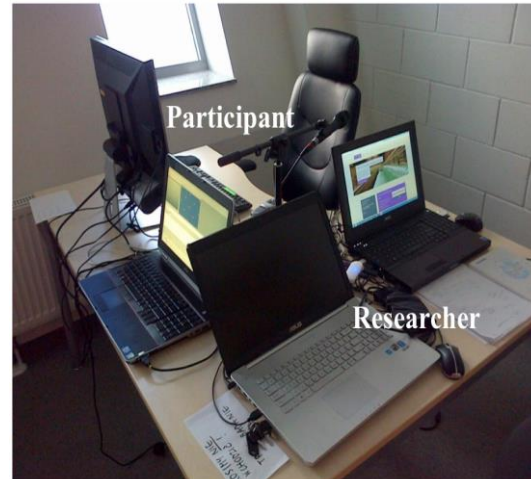
Figure 4.9 Researcher's laptop   **Figure 4.10** Participant's station and eye tracker
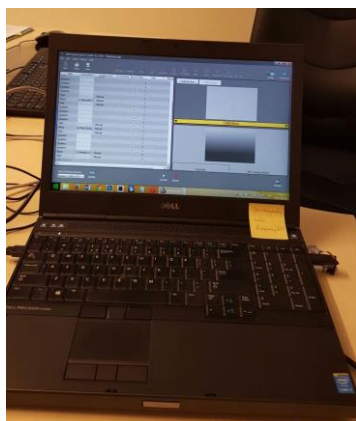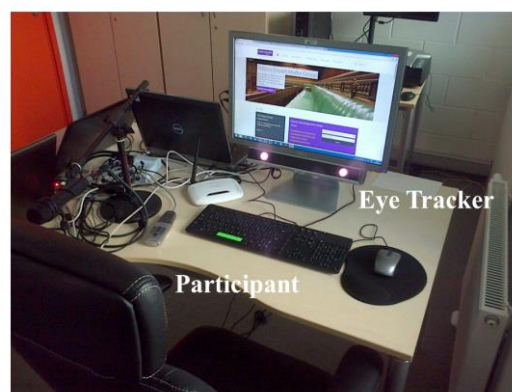
It was also ensured there was a constant ambient light throughout the duration of all eye-tracking sessions and windows were covered with blinds to avoid the sun light entering the lab, which might have interfered with the infrared red light flashing from the eye tracker.

All participants were scheduled for a specific day and time. I allocated up to 90 minutes (eye-tracking component lasted 20 minutes) per person to complete the eye-tracking session. Participants were offered a break in the meantime, if they wished, yet it was scarcely necessary as a great majority of respondents finished the session within 60 minutes. Prior to the start of the recording, they were asked to sign the consent form and complete the self-assessment questionnaire regarding their English proficiency and assess their listening comprehension according to the one of the language levels (i.e., C2/C1, B2/B1, A2/A1) as recognised by CEFL (Appendix 1). On the basis of the questionnaire, between 6 and 8 subjects were allocated to a subgroup of high, medium and low level of English proficiency in SUB_VO and VO_SUB (see subsection 4.6.3) before the recording started, as shown in Table 4.10 below.

Table 4.10 The number of participants allocated to each subgroup in SUB_VO and VO_SUB groups according to their level of proficiency in English

Groups	Group SUB_VO	Group VO_SUB
HLPs	7	8
MLPs	8	7
LLPs	8	6

Subsequently, the functionality of eye tracker was explained, how the session would unfold and how long it would last. Each participant was asked to take a seat on a black rotating chair. Depending the height of a respondent, the chair and the position of eye tracker was adjusted to make sure his/her eyes are located in the centre of the screen. Each participant was positioned very accurately, that is, at approx. 60 cm distance in from of eye tracker, as shown in Figure 4.11.

Figure 4.11 A participant seating in front of the monitor with eye tracking in the AV Lab at UW
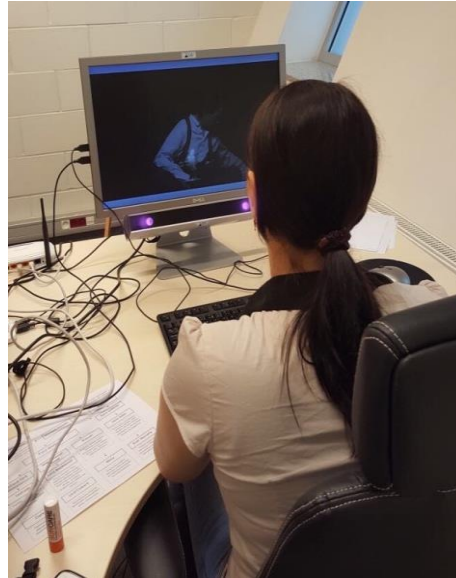
Once the participant was seated in front of the eye tracker in the correct position, the calibration process started. Since my experiment did not involve pure text reading but watching clips with subtitles, a standard 5-point calibration was used in which each participant had to follow with his/her eyes the points on the screen known as calibrate dots (see Figure 4.12). During this process, the eye tracker measures the characteristics of the eye including shape, refractions and reflections in different parts of the eye (SMI, 2016). In order to obtain maximally precise, data the value of X and Y should be below 1.0, otherwise it is recommended to recalibrate the participant. The calibration process is followed by the validation procedure to verify the outcome of calibration. Calibration and validation (Figure 4.13) should be repeated until the required precision is achieved.

Figure 4.12 Calibration procedure in SMI Experiment Center

Figure 4.13 Calibration values in pictures 1 (poor calibration quality) and 2 (good calibration quality)
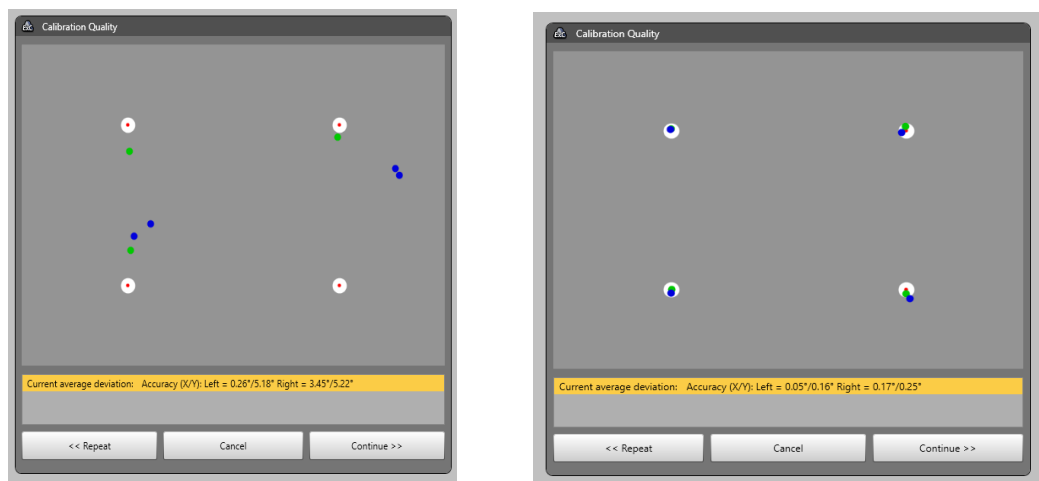
When calibration and validation were approved, the participants were assigned to one of the language groups (high, medium, low) and were asked to watch one set of videos (SUB_VO or VO_SUB). 2 sets of videos were imported in AVI format with a 1920×1080 resolution and a fit-to-the-screen option to Experiment Center prior to recording. As mentioned in subsection 4.7.3, 4 participants did not achieve sufficient calibration values during the calibration phase, and as a result they were excluded from the eye-tracking recording and data analysis.

The experiment started with the welcome page and instructions for the participants about the study. The next slide outlined the contextual introduction to the first scene and the clip was displayed in the following slide. The participants were asked to answer four questions after each video. Figure 4.14 illustrates the order of all elements included in the experiment design.

Figure 4.14 Sequence of the presented tasks in SMI Experiment Center
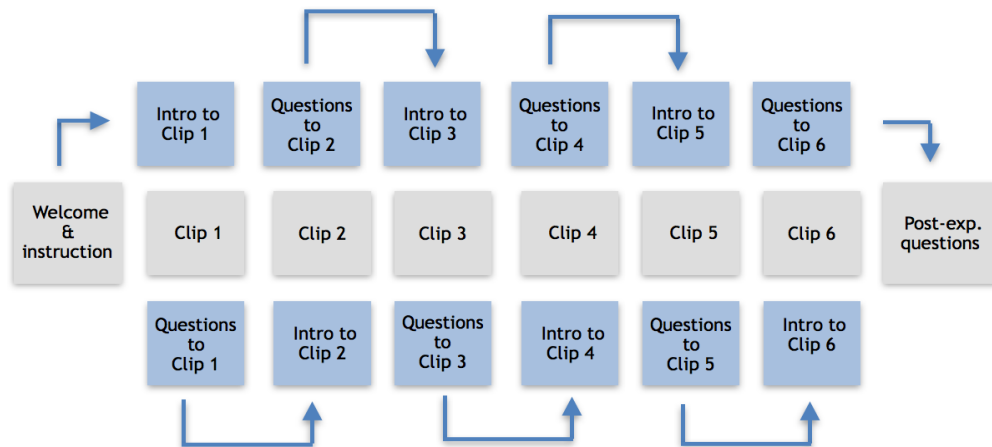
The eye-tracking experiment design was displayed in SMI Experiment Center software (available at the University of Warsaw), as presented Figure 4.15.

Figure 4.15 Experiment design in SMI Experiment Center

Type	Source/Name	Duration [ms]	Fit to Screen	Record Data	Task
Calibration					
Validation				×	
Text	RichText.rtf	Manual		☐	SUB_VO
Movie	1_Cemetery_S	Manual	×	×	VO_SUB
Movie	1_Cemetery_V	Manual	×	×	
Question				☐	
Text	RichText1.rtf	Manual		×	SUB_VO
Movie	2_Watson_SUE	Manual	×	×	VO_SUB
Movie	2_Watson_VO.	Manual	×	×	
Question				☐	
Text	RichText2.rtf	Manual		×	
Movie	3_Cawford_SU	Manual	×	×	SUB_VO
Movie	3_Cawford_VO	Manual	×	×	VO_SUB
Question				☐	
Text	RichText3.rtf	Manual		×	
Movie	4_Jungle_SUB.	Manual	×	×	SUB_VO
Movie	4_Jungle_VO.a	Manual	×	×	VO_SUB
Question				☐	
Text	RichText4.rtf	Manual		×	
Movie	5_Stag party V	Manual	×	×	SUB_VO
Movie	5_Stag Party_S	Manual	×	×	VO_SUB
Question				☐	
Text	RichText5.rtf	Manual		×	
Movie	6_Train_SUB1.	Manual	×	×	SUB_VO
Movie	6_Train_VO.av	Manual	×	×	VO_SUB
Question				☐	
Text	RichText6.rtf	Manual		×	
Question				☐	
Text	RichText7.rtf	Manual		×	
Question				☐	
Question				☐	
Question				☐	
Question				☐	
Text	RichText8.rtf	Manual		×	

As shown in Figure 4.15, the order and duration of the stimuli presentation could be adjusted in the Experiment Center. In the present study, the viewers controlled their own timing when

answering the questionnaire as additional time pressure might have influenced their responses. Therefore, they used either the keyboard or the mouse to move along the experiment. Throughout the duration of the experiment, a real-time observation of the participants' eye movement was made to ensure that they come back to their calibrated position while watching the stimuli, as shown in Figure 4.16 below.

Figure 4.16 Monitoring of participant's eye movements during the eye-tracking session
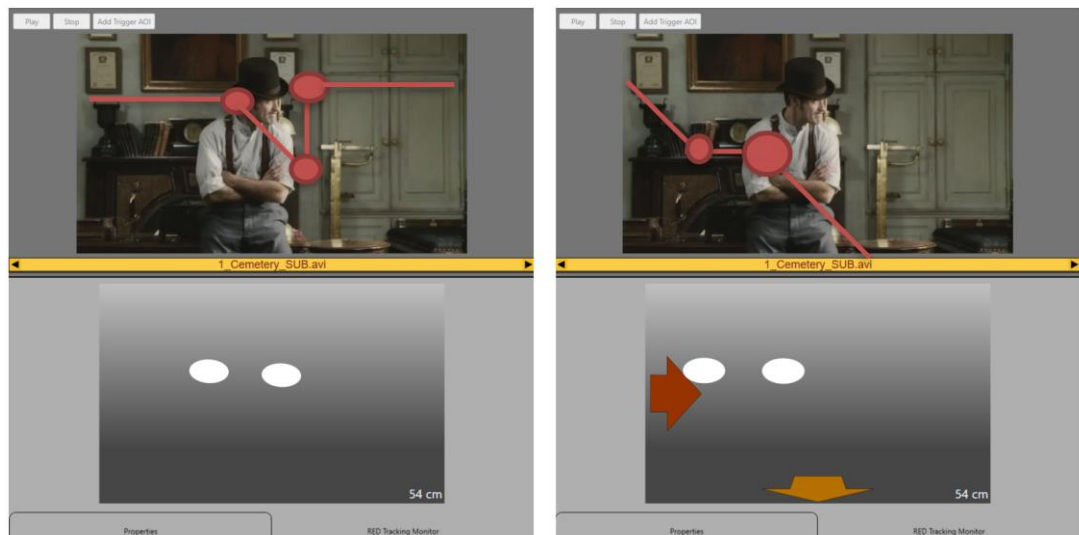
Although my presence might have had an impact on the participant's behaviour during the experiment, I stayed in the lab throughout the duration of each session with each participant in case some technical problems came up and to ensure that they come back to the calibrated position before watching the next video. None of the participants reported to me that he/she felt uncomfortable in any way due to my presence during the eye-tracking session. Once the recording was completed, they were asked to answer a few more questions in the post-experiment questionnaire and were offered a cup of coffee and a piece of cake as a thank you. The volunteers were not offered any monetary reward for their participation in the experiment.

Once all recording sessions were completed, the quality of eye-tracking data necessary to run the analysis was inspected exploring gaze time on screen. The gaze time proved to be the most reliable method in regard to reception studies of multimodal texts in AVT (Orrego-Carmona, 2015). The percentage of gaze data was extracted from SMI BeGaze software. To decide whether the participants should remain in the data set, I adopted O'Brien's (2010) threshold that helped me flag these participants whose gaze time on screen

was 70% and above of the total time they spent looking at the screen. On this basis, the data of all participants were accepted in the experiment.

Subsequently, the eye-tracking data was exported using the event statistics in SMI BeGaze analysis software for quantitative analysis and statistical tests that are frequently applied when analyzing eye-tracking data. Despite the small sample size in each subgroup of SUB_VO and VO_SUB in the current experiment (see Table 4.9), I still decided to include statistical tests and interpret the data using descriptive and inferential statistics. Additionally, the visualisations were exported as AVI file (e.g., heat maps, scan paths) to see how data was presented visually in all language levels subgroups of SUB_VO and VO_SUB and to complement the quantitative calculations, if necessary. SMI BeGaze software also enabled the export of the questionnaire questions with the participants' responses.

4.8 Conclusions

The main goal of Chapter 4 was to present the proposed methodological apparatus that was designed to investigate how multimodal irony is construed, relayed and grasped by all viewers in the subtitled and voiced-over clips of SH1 and SH2. The experimental design was also supposed to facilitate the comparative analysis of how the participants representing different levels of language proficiency (high, medium, low) consume information through verbal and non-verbal modes in the selected scenes. In order to increase the reliability and validity of the experimental design, the multimodal transcription, eye-tracking technology and questionnaire have been integrated for the purposes of triangulation.

With the adapted model of multimodal transcription, it was explored how multimodal irony is construed in the two Sherlock Holmes films under analysis. This process allowed me to specify how verbal and non-verbal elements co-deploy to construe meaning, and more importantly ironic meaning on screen. Since the multimodal transcription is largely an intuitive process, other methodological tools had to be incorporated to produce more objective data. As a result, eye-tracking technology was added to the methodological apparatus.

Eye tracking is particularly valuable in the context of the current study as it enabled me to detect what strategies were applied in order to make sense of the presented stimuli and to determine how the viewers at different language levels consume the multimodal text under scrutiny. More importantly, eye tracking helped me understand and compare how they share their attention throughout the different parts of the screen. It is thus possible to demonstrate the discrepancies in the viewing behaviour among different participants when watching the selected clips with subtitles and voice-over.

Although eye tracking is a very powerful technology, it cannot reveal what people think when they are exposed to a given stimuli. In order to get a more in-depth insight of the participants' thinking processes, a questionnaire has been incorporated into the methodological design to probe the comprehension of the intended ironic meaning. The coding apparatus was devised to quantify the participants' responses to the open-ended questions. The inclusion of the questionnaire thus allows me to support the eye-tracking data and correlate them with the answers to the selected scenes featuring the use of multimodal irony.

Considering the mix-method approach as well as the triangulation of three different methods, it is argued that the proposed methodological apparatus represents a comprehensive mechanism to explore the phenomenon under study and generate reliable results. In the next chapter, the outcome of the multimodal transcription is presented in which it has been examined how meaning, and particularly ironic meaning is construed and conveyed into Polish with subtitles and voice-over.

5 ANALYSIS AND DISCUSSION OF DESCRIPTIVE DATA

5.1 Introduction

Chapter 5 constitutes the first out of three analytical parts of this projects. It seeks to analyse the construal of multimodal irony in the film dialogue of the selected 6 clips of SH1 and SH2 and compare the way ironic meaning is relayed with subtitles and voice-over in Polish using the adapted model of Baldry and Thibault's (2006) multimodal transcription (see Section 4.4). The multimodal analysis has demonstrated how the verbal and non-verbal modes intersect to produce and convey ironic meaning within the multimodal environment of the subtitled and voiced-over clips in the light of Sperber and Wilson's (1992) echoic theory of irony. As an illustration, the tables with the analysis of multimodal irony for clip SH1_1 have been presented in the body of this chapter, while the multimodal tables for the remaining clips have been included in Appendix 5.

This chapter is structured as follows. Section 5.2 explores the construal of multimodal irony in the original dialogue of the selected scenes under scrutiny. Particular emphasis is placed upon the verbal and non-verbal modes of film communication and the way they co-deploy to produce the intended ironic meaning. Sections 5.3 and 5.4 examine and contrast the English dialogue *vis-à-vis* the subtitled and voiced-over clips featuring the use of multimodal irony. Section 5.5 further debates whether and/or to what extent the two modalities of audiovisual translation, namely subtitling and voice-over, and the level of English-language competence impede the reception of multimodal irony on screen in the context of this study. Finally, this chapter concludes with Section 5.7 and summarises the contribution of the verbal and non-verbal modes in the generation of multimodal irony in the present study.

5.2 Construal of Multimodal Irony in the Source Texts

As illustrated in Section 3.5, the film directors use dialogue to expand the film text's "ironic capabilities" (Kozloff, 2000: 54) in the context of the current project. Through the narrative and aesthetic functions, the filmmakers enable audiences to uncover their communicative intentions, and thereby, come to the conclusion that the echoic quality of the utterances and the speaker's attitude of dissociation towards this utterance is meant to be interpreted as ironic (Sperber and Wilson, 1992; Yus, 2000). Without a doubt, dialogue is essential in the generation of ironic meaning in the film text, but it is not capable of informing the construal
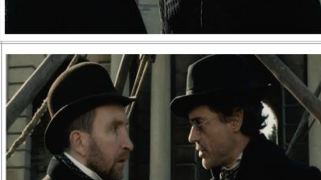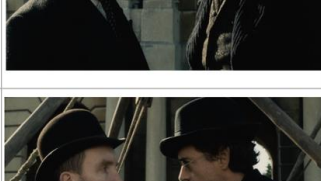
and functions of multimodal irony in the clips under study independently of the other non-verbal semiotic resources of film language. This section outlines how the verbal mode (i.e., dialogue) in collaboration with the non-verbal modes pertaining to mise-en-scène, cinematography, editing and sound jointly construe multimodal irony in the two layers of film communication, namely, at the vertical and the horizontal layer (e.g., Kozloff, 2000: 170-200). Additionally, it is explored how these meaning-making semiotic resources let multimodal irony fulfil its humorous functions in the selected clips from the two Sherlock Holmes films. The analysis of 6 clips from SH1 and SH2 has been presented in the following pages. The instances of multimodal irony have been illustrated with some concrete examples underscoring the importance of the verbal and non-verbal modes in its production to further corroborate this claim.

Let us consider the first clip **SH1_1** in which a wide array of semiotic resources plays a pivotal role in the construal of multimodal irony. In this scene, Sherlock investigates the process of bringing up the coffin in which the villain Lord Blackwood is buried. Given that the investigation goes at a very slow pace, Sherlock asks the Scotland Yard's Inspector: "Lestrade, what's with the coffin?" and Lestrade responds: "We are in the process of bringing it up now". Sherlock continues the conversation saying "I see" (frame 1) "Hmm. All right" (frame 3) while taking a glance at Lestrade's policemen who stand still and do not seem to be interested in completing the task entrusted to them (frame 2). Sherlock turns back to Lestrade and asks him again (frames 4-5): "At what stage of the process? Contemplative?" but Lestrade fails to identify the echoic nature of Sherlock's utterance and his disapproving attitude towards him. In this fragment, the intended ironic meaning is enlivened by the use of sarcasm when Sherlock is deliberately insulting him with intention to ridicule the activities undertaken by Lestrade and his troops, which under current circumstances are not very useful as these will not help to find out whether Lord Blackwood is actually inside the coffin or not. According Garrand (2000) sarcasm also serves the creation of humour in film dialogue. In this instance, apart from Sherlock's dissociative attitude, the viewers' smile is also bound to be evoked by a feeling of superiority, which, as believed by Vandaele (1999/2002), is used for humorous purposes. In this fragment, Lestrade does not realise that he is being mocked by Sherlock but viewers watching this scene have access to more information, and therefore can feel superior to Lestrade as the target of mockery. The acoustic modes are also of paramount importance in this instance of multimodal irony. Sherlock's voice modulation is here instrumental to the expression of his attitude of dissociation towards the proposition echoed to mock Lestrade and his troops, as shown in

the multimodal analysis (Figure 5.1). Particularly, Sherlock's flat and raising intonation, heavy stress, low pitch, slow speech rate and pauses actively participate in the meaning making process in this clip, in addition to the film score, as it is also the case in the previous studies (e.g., Attardo et al., 2003; Kreuz and Roberts, 1995). Apart from the acoustic modes, the non-verbal kinesic and visual modes (i.e., mise-en-scène, cinematography and editing) are also highly significant in the production of the intended meaning in this clip. For instance, Sherlock's deadpan face directed straight into Lestrade's face in concert with his head movements and some of his props (like sunglasses) underline the disapproving character of his attitude towards the Inspector in medium close-up shots (frames 3-5). In addition, crosscutting editing technique enables the audience to see Lestrade's policemen relaxing instead of working (frame 2) practically at the same time to "highlight an ironic state of affairs" (Macdowell, 2016: 117). The co-deployment of verbal and non-verbal resources has been presented in Figure 5.1 below.

Table 5.1 Multimodal transcription of clip SH1_1 frame by frame to distinguish verbal and non-verbal resources that contributed to the construal of multimodal irony

Clip **SH1_2** is another illustration of a combination of the verbal and non-verbal modes as means of creating multimodal irony. In this scene, Sherlock spends a night using black magic rituals to find out how the villain Lord Blackwood plans to fulfil his next prophet and murder another person. Sherlock wakes up the next day in the morning and realises that Irene and Watson came to him with a visit. Watson sits in a chair and observes Sherlock coming back to life. When the detective is trying to sit on a bed, Watson comments on his looks: "You look gorgeous" (frames 2-5). In this fragment, the intended meaning is construed with Watson's ironic riposte singling the incongruity between Watson's comment and the way Sherlock is presented in the given context. It is thus blatantly visible that the filmmakers did not intend to convey the literal meaning of Watson's statement in this scene the word "gorgeous" is the opposite of what he really intends to say in order to mock Sherlock's hideous appearance. In this fragment, film dialogue conveys vocabulary choices that are typically used to trigger ironic interpretations – and that Yus labelled linguistic cues (Yus, 2002). Although some of these expressions may not be irony-specific, under certain contextual circumstances, they indicate the speaker's dissociative attitude. Seto (1998: 246f) gives an account of linguistic cues of a stylistic, syntactic or lexical nature which may help the audience to interpret ironic meaning. For instance, on the lexical level "single words intrinsically charged with a very high positive meaning" can contribute to the construal of irony (Yus, 2002: 11). It is also the case with the word "gorgeous" which is loaded with a positive meaning but is not intended to be understood as Watson's attempt to be polite towards Sherlock but to mock the way he looks. Here, humous arises when the viewers spot this incongruity. However, it is only possible if Watson's ironic riposte is accompanied by the visual and kinesic modes that in this clip are highly significant, as illustrated in the multimodal transcription (Appendix 5 SH1_2). For instance, within the visual mode, there are non-verbal semiotic resources such as props, pieces of garment and elements of a setting that are involved in the meaning making process in this sequence. In regard to settings, Sherlock wakes up in a dirty and unpleasant room with a few pieces of old furniture (frame 1). There are burned candles and painted anagram on the floor. In addition, Sherlock looks tired and half asleep dressed in his dirty and untidy clothes. It seems that Sherlock was working very intense last night. Watson, on the other hand, is properly dressed and is sitting relaxed in font of Sherlock. Humorous irony arises when viewers spot the incongruity between what they expect to see on the basis of Watson's comment and what they actually see in the given context (Sherlock who looks rather terrible but not gorgeous). While mise-en-scène excels in the expression of Watson's dissociative attitude irony, it cannot co-create

multimodal irony without the engagement of cinematography. Here, camera functions come into play by visualising the distance and angle between Watson and Sherlock (frame 1). Through the use of a low camera angle, Watson is looking at Sherlock from above which crease a feeling of superior over Sherlock. In addition, continuity editing and the establishing shot also relay irony visually. The establishing shot allows the audience to gain some contextual information in a medium long shot (frame 1) first and then spot Watson's kinesic behaviour while uttering his comment (frame 2) in a medium close-up shot, as Table 5.1 illustrates:

Table 5.2 Illustration of the establishing shot in SH1_2

Frame No	Time (h:m:s)	Frame	Visual Image	Kinesic Action	Sound
1	1:27:33		CP: stationary (vertical) CD: MLS, HA VC: settings, old room, sign of the floor, props, costumes ED: continuous VF: Sherlock wakes up. Watson looks at Sherlock from above.	Watson sits straight and is looking at Sherlock from above. Sherlock slowly sits on the bed.	♫
2	1:27:35	Wyglądasz cudownie.	CP: stationary (horizontal) CD: MCS, SA VC: settings, old room, sign of the floor, props, costumes ED: continuous VF: Watson watches Sherlock as he wakes up.	Watson sits straight on a chair. He raises his head looking at Sherlock from above squinting his eyes. Watson smiles and Sherlock with pity. ☺ (><)	WAT: You look gorgeous. ♫ p, M, LP, #

In this instance, the kinesic resource, that is, Watson's body language serves as an explicit marker of irony (Yus, 2002) which makes his attitude of dissociation even easier to detect. Watson sits up straight in a nonchalant position, holds his shoulders back and keeps his head high looking at Sherlock from above with squinting eyes and a gentle smirk (frame 2-5). Additionally, the acoustic modes support the multimodal construal of irony through Watson's particular vocal patterns, including the slow speaking rate, flat intonation and stress on the word "gorgeous".

Apart from ironic ripostes, the intended ironic meaning is also construed in SH1 and SH2 through "a contest of wit" which Abrams (1993: 220) refers to as repartee which is achieved with the help of verbal and non-verbal modes. Consider clip **SH1_3** in which Sherlock is brought by Inspector Lestrade to the chamber of Lord Coward - the collaborator of the villain Lord Blackwood. Sherlock starts teasing with the Lord and hides behind a wall of smoke, while Coward is walking around the room and is trying to shoot him. As Sherlock aims to find out when Lord Coward intends to commit his next murder, he starts playing with Coward in ironic repartee asking a number of ambiguous questions to deliberately mock Lord Coward's presumptuousness and self-confidence. Finally, he reveals the most important information and says to Sherlock: "We are starting at noon". At this moment, the camera shot shows Sherlock who is sitting relaxed in a chair smoking a pipe and does not really mind that Lord Coward has already attempted to shoot him several times. In this moment Sherlock is expressing his attitude of disapproval and responds to Coward: "There isn't any time to waste then. Is there?" (frames 1-3). The intended ironic meaning is thus created not only with Sherlock's "excessive playfulness" (Bordwell, 2006) and "contest of wit" (Abrams, 1993) but also it is signaled by the help of a question tag "Is there" and Sherlock's intonation is additionally highlighting his ironic utterance. Thus, the acoustic modes are here vital in the construal of multimodal irony as Sherlock's slow speech rate, low pitch and flat intonation highlight his mocking attitude towards Coward while producing his utterance (see Appendix 5 SH1_3). More importantly, however, the characteristic quirky instrumental music is accompanying Sherlock and Lord Coward in their ironic repartee and is intensifying the hilarity in the situation. The film score consists of rather unconventional instruments including a banjo, a cimbalom, a squeaky violin, an accordion, dumbeks, a "broken pub piano" or the experibass. The quirky music underscores Sherlock's mocking attitude towards Lord Coward, particularly in the sequence shots in which Lord Coward attempts to shoot Sherlock – who is comfortably sitting on the other side of the room smoking his pipe (frame 1) and then delivers his utterance (frame 2-4). The co-deployment of diegetic and non-diegetic sounds in this scene intersect with camera distance (a medium long and medium close-up shot) and establishing shot. In medium close-up shots, Sherlock's facial features are particularly discernible such as wide-open eyes, raised eyebrows, "a blank face" that contribute to the creation of irony in a multimodal way, as it is also the case in the previous research of Attardo et al. (2003).

In clip **SH2_1** the verbal mode also does not produce the intended ironic meaning in isolation from the non-verbal meaning-making semiotic resources. In this scene, Watson

pays a visit to Sherlock after a longer absence. He opens the door to Holmes' flat, sits in a chair and starts reading a newspaper, while waiting for his friend to leave his hideout. Suddenly, Sherlock appears in front of Watson in his camouflage. Knowing that they celebrate his stag night tonight, Watson expects his friend to be dressed properly. After a few seconds, Watson takes a glance at his costume and comments: "I'm not going out with you dressed like that". In response, Sherlock is expressing his attitude of dissociation towards the proposition echoed to ridicule Watson's outfit too and says: "Would you prefer I joined you in the fashion faux pas of wearing fine military dress with that heinous handmade scarf clearly one of your fiancée's early efforts?" (frames 1-10). In this clip, Sherlock is uttering a number of personal insults against Watson mocking his dress code in a wry and sarcastic manner. Sherlock's cut wit is also intended to attack Watson's marriage with his future wife Mary. These "cerebral tennis matches" between Sherlock and Watson (Lyall, 2009) are a known illustration of the special ironic and humorous character of their relationship and friendship. Moreover, Sherlock's witty and sarcastic banter in combination with his body movements and gestures becomes even more amusing for the audience, as shown in the multimodal tables (Appendix 5 SH2_1). In this fragment, multimodal irony featuring some comedic effects is thus composed with a combination of the visual (cinematography, editing and mise-en-scène) and kinesic modes (mise-en-scène). The first few medium long shots (frames 1-3) present Sherlock from a low angle when he is disrespectfully waving with a piece of cloth in front of Watson which reinforces "in the detective a certain air of arrogant superiority" over his companion (Tallon and Baggett, 2012: 117). Following medium close-up shots (frames 4-10) show Sherlock at closer distance and allow the audience to spot his facial expressions in ample detail. A medium close-up shot also gives prominence to other non-verbal semiotic resources pertaining to mise-en-scène (body language, costumes, props, eye contact) that contribute to the construal of the intended meaning in this excerpt. The acoustic mode including Sherlock's flat intonation and low pitch constitute a supportive source of information in the composition of multimodal irony in this scene.

In clip **SH2_2** the non-verbal modes assist the verbal mode with the generation of ironic meaning. In this scene, Sherlock invites Watson to a restaurant to celebrate his stag night. They take a sit at a table and Sherlocks orders champagne to drink his health. Watson expects that his colleagues from the military or medical school will join them for the celebration. After a while Watson realises that his accountancies are not going to show up because Sherlock forgot to invite them. When Watson reproaches Sherlock for ruining his

stag night, Mycroft (Sherlock's brother) comes to their table and says: (frame 1-5): "Your very good health, doctor. Shame none of your friends could make it." In this fragment, Mycroft's mockery is expressed through the literal meaning, although his ironic riposte is not intended to be understood literally but ironically. In truth, Mycroft does not feel sorry for Watson that Sherlock is a terrible best man but is mocking him in a humorous way. In this clip, Mycroft's ironic comments could be easily misunderstood, if his attitude of disapproval was not accompanied by a combination of the acoustic, kinesic and visual modes of film communication. As shown in the multimodal analysis of the descriptive strand of this study, a composition of diegetic and non-diegetic sounds significantly contributes to the interpretation of Mycroft's dissociative attitude as ironic (Appendix 5 SH2_2). When Mycroft utters an ironic comment with a raising intonation, Sherlock and Watson exchange looks and then burst out laughing in a wry manner (frame 6-10). In this scene, this form of artificial laughter, and more importantly "antiphonal laughter" (Bryant, 2011) is the most obvious and powerful marker of ironic meaning. Although multimodal irony is predominantly construed via the non-verbal acoustic semiotic resources, Mycroft's ironic attitude is additionally manifested by the visual and kinetic modes. Particularly, Mycroft's facial expressions like squinted gaze, raised eyebrows, wry smiling and smirking in unison with his body language including head movements, gestures and gaze direction (Mycroft is looking at Watson from above) intersect with the non-verbal acoustic cues as supportive sources of information. In addition, medium long and medium close-up shots enable the viewers to perceive the characteristics of Mycroft's kinesic behaviour, while editing shots quickly change to show Sherlock, Watson and Mycroft while they are engaging in the artificial wry laugh.

Clip **SH2_3** is another good case in point where the verbal mode is supported by an amalgamation of the kinesic, visual and acoustic modes in the construal of multimodal irony. In this fragment, Watson and Mary are going by train for a honeymoon. While they start celebrating, Sherlock enters their carriage dressed like a woman. It turns out that Watson and Mary are in danger as the troops of the villain Professor Moriarty intend to shoot them. Sherlock builds a line of attack to defend them and count on Watson's military skills to shoot a soldier in the exact moment he planned and gives him the opportunity to kill one of the soldiers operating the machine gun. Unfortunately, Watson only grazes the soldier, but does not kill him. Sherlock comments on Watson's failure: "I said make it count. How many windows must I provide?" mocking Watson's military background. In this scene, Sherlock is asking a wry and rhetorical question to express his attitude of disapproval to sneer at

Watson's shooting capabilities. In this instance of multimodal irony, the filmmakers employed ironic metaphor to produce the intended ironic meaning via the verbal mode. In the sequence, according Sherlock's ironic dialogue, he gives Watson "a window of opportunity" to shoot the soldier. By the idiomatic expression "a window of opportunity" Sherlock means to indicate that he gives Watson a chance to shoot accurately. When Watson fails and Sherlock wonders how many other chances Watson needs to defend them under current circumstances due to attack. Sherlock's dissociative attitude is additionally expressed by the non-verbal semiotic resources pertaining to cinematography and editing in tandem with mise-en-scène and sound, as illustrated in the multimodal tables (Appendix 5 SH2_3). A pace of action is here very fast and dynamic. Fast changes from one shot to another – in the form of fast, frequent and abrupt cutting – are used principally to build tension, suspense and excitement. As Kellison (2012: 155) confirms, "an editor can start with longer cuts, then more frequent cuts that surprise the viewer or built suspense". Editing is immensely influenced by sound, and in particular by the movie's rhythm (Chandler, 2009). Similarly, to music, the rhythm of editing is highly significant in films; in cases like SH2_3, it results in fast, slow sequences or abrupt and dramatic shots affecting the audience's perception of the scene (Chandler, 2009) and contributing to the creation of a sense of excitement, and comic effects in this scene. In SH2_3 this type of editing enriched with camera position, Watson and Sherlock's facial expressions and body language and Sherlock's ironic comments thus aims to combine suspense with a humorous dimension of the scene leading to the composition of irony multimodally. Additionally, sound effects, Sherlock's voice modulation in concert with the film score incite and intensify emotions, feelings and Sherlock's attitude of dissociation contributing to the creation of this instance of humorous irony in the given scene.

In the above discussion, it has been demonstrated with the help of the multimodal apparatus how the verbal and non-verbal modes of film language co-deploy to compose ironic meaning on screen. The analysis has also revealed whether the manifestation of the on-screen character's dissociative attitude has been supported by the information relayed acoustically or visually. It is also worth highlighting that the verbal mode (film dialogue) is not capable of producing ironic meaning alone but in collaboration with cinematography and editing in tandem with mise-en-scène and sound. Otherwise, the filmmakers would run the risk that the intended meaning could be understood as literal and the ironic effect could be lost. In the next section, it has been analysed how the audiovisual translator rendered

multimodal irony as it was intended by the filmmakers in the original dialogue with subtitles and voice-over in the TL.

5.3 Relay of Multimodal Irony in the Subtitled Texts

In Section 2.4, some challenges related to the translation of ironic meaning with subtitles have been outlined. According to the audiovisual translation scholars, the contribution of the non-verbal visual and acoustic cues is highly significant in the transfer of irony in audiovisual productions (Pelsmaekers and Van Besien, 2002; Zabalbeascoa, 2003). In this section, it is examined how irony is rendered in the subtitles - the most popular modality of audiovisual translation in the global audiovisual industry these days (Díaz Cintas, 2008) highlighting the importance of the non-verbal semiotic resources in this process. Drawing on the multimodal analysis of the descriptive strand, two broad categories of irony relay in the film texts have been put forward: *preservation* (i.e., the wording and the form of the source language dialogue have been preserved in the target language subtitles while rendering ironic utterances) and *modification* (i.e., the wording and the form of the source language dialogue have been modified in the target language subtitles while rendering ironic utterances. The analysis showed that 6 instances of multimodal irony identified in SH1 and SH2 and included in the present descriptive data set were handled by the audiovisual translator in terms of preservation and modification. That is to say, the subtitler opted for the use of modification to a greater extent (4 out of 6 clips) than for preservation (2 out of 6 clips) in both Sherlock Holmes films when transferring multimodal irony into subtitles. Only in two clips (SH1_2 and SH2_2) the subtitles rendered on the lexical level as they were composed in the original dialogue. In the subsections that follow, the terms preservation and modification are explained and illustrated. It is also debated whether and/or to what extent the ironic effect has been retained or totally lost when using the categories of translation to relay multimodal irony in the subtitled version of the film text under analysis.

5.3.1 Preservation of multimodal irony in the subtitled texts

Preservation is here defined as "the intact transference" of the SL ironic content into the TL (Desilla, 2009: 209). It means that the subtitlers render each word of the SL wording that the filmmakers communicated in the original dialogue. In general, verbatim translation is only occasionally possible due to (i) conventions around the limited space and time on screen (see Díaz Cintas and Remael, 2007; Georgakopoulou, 2010) and (ii) linguistic and cultural differences between the source and the target languages and cultures. Given that Polish and

English are typologically very different languages (Ożańska-Ponikwia, 2013) and that irony is a complex linguistic phenomenon composed in the cultural context, the transference of ironic content appears to be challenging. The preservation of the original wording in the subtitles does not necessarily ensure that the audience identifies the echoic and ironic quality of the utterance, and thereby the speaker's attitude of dissociation towards the proposition echoed. In truth, there are instances of multimodal irony in which the subtitles followed closely the form of the source language. Nevertheless, viewers would not retrieve the ironic meaning of the utterance, if no access to the non-verbal modes of film communication was given. In the following pages, it has been analysed whether the translation of the source dialogues with this category retains the contribution in the TL that the verbal mode made to the instances of multimodal irony in the SL, as described in Section 4.2, and gauge the extent to which the ironic effect is retained in the subtitles.

SH1_2 is an example of ironic utterance that is rendered verbatim in the subtitles. That is, Watson's dialogue is preserved by the subtitler as "Wygladasz cudownie" (BT: "You look gorgeous"). In this scene, the target translation presumably triggers the same understanding of Watson's ironic riposte in the mind of the target audiences. The tones of Watson's mockery expressed by his ironic attitude towards Sherlock are thus relayed in the subtitles, too. In this instance, multimodal irony should be processed together with the visual and kinesic modes by the source and target viewers. Otherwise, the viewers may have difficulties to interpret Watson's utterance as ironic and grasp the literal or positive meaning of his statement. In clip **SH2_2**, the contribution of the non-verbal modes of film communication is also essential in the perception of Mycroft's dissociative attitude towards Watson. In this scene, each word delivered by Mycroft in the original dialogue is also closely rendered into the subtitles as "Szkoda, że przyjaciele nie przyszli" (BT: "Shame your friends did not come") presenting the way ironic utterance was construed by the filmmaker's in this fragment. In the absence of the non-verbal acoustic semiotic resources, Mycroft's dialogue can be easily misunderstood. Moreover, a low camera angle in concert with Mycroft's smirk and Sherlock's jeering laugh, among others, are expected to help the audience identify Mycroft's dissociative attitude towards Watson, and thereby retrieve ironic meaning in the subtitled text. As a result, it is assumed that the target audiences will are able to retrieve the intended ironic meaning as the source audiences with the help of verbal and non-verbal semiotic resources.

5.3.2 Modification of multimodal irony in the subtitled texts

In the context of this study, modification is termed as the transformation of the SL utterance into the TL utterance. It means that the source language dialogue undergoes several changes at a textual or sentence level before it is rendered in the form of paraphrase, condensation or reduction. That is to say, some words or phrases including repetitions, vocatives or intensifiers, to name just a few, are deleted or paraphrased in order to convey the meaning of each expression or sentence using words that sound natural for the target viewers without rendering each word of the original dialogue. In interlingual subtitles, text modification in the form of reduction and condensation is a widely used practice predominantly due to the limited space and time available to display subtitles at the bottom of the screen (see Díaz Cintas and Remael, 2007; Georgakopoulou, 2010; Tomaszkiewicz, 2006).

So far modification has been recognised as the most frequent way of relaying the utterances containing ironic content into subtitles (Pelsmaekers and Van Besien, 2002). A number of lexical items such as "forms of address, hedging and intensifying devices, word choice, cultural references, markers of prosodic features, repetitions and echoes" are most likely paraphrased, condensed or omitted in subtitles (Pelsmaekers and Van Besien, 2002: 253). As shown in the analysis, the subtitler modified the content of the source dialogue in 4 out of 6 instances of multimodal irony in SH1 and SH2. In the case of the category of modification, the subtitler had the opportunity to "tamper" with the transfer of ironic meaning in the TL that was originally intended by the filmmakers in the source texts. The following part of this section moves on to exemplify the changes made to the original dialogues and outline in greater detail whether these alterations made to the instances of multimodal irony in the SL have the impact on the transfer of the ironic effect in the TL subtitles.

Let us consider **SH1_3** as an illustration of the modified subtitles in the TL. In this clip, Sherlock's response to the proposition echoed by Lord Coward is reduced and subtitled into Polish as "A więc nie ma czasu do stracenia" (BT: "So there is no time to lose"). The question tag "Is there?" in the original dialogue is ancillary for irony relay since Sherlock is not really asking a question; instead, he is mocking Lord Coward and expressing his dissociative attitude towards him. However, the subtitler omitted the question tag in the TL version. A possible reason underpinning the omission of the question tag can be linked to the fact that these forms of "short answers and reply questions are not as frequent in Polish as they are in English, so Poles tend to avoid them altogether" (Swan and Smith, 2001: 168). According to Wierzbicka (2003: 38) there are only six types of questions tags used in Polish,

e.g., prawda? (BT: true), nice? (BT: No?), co? (BT: what?), to name just a few, that are not tense related and their repetitive use would make Polish language sound grotesquely. Nevertheless, whether the ironic effect is (or not) diminished in the TL version due to the omission of the question tag in the subtitles has to be tested empirically.

Clip **SH1_1** is also a good case in point in which the source text is reduced and paraphrased in the target text. In this scene Sherlock's sarcastic dialogue is subtitled into Polish as: "Na jakim etapie? Zastanawiania się?" (BT: "At what stage? Thinking about?"). Although the word "process" was omitted and "contemplative" was transferred with a synonymous expression "thinking about" in the subtitles, this alteration does not seem to interfere with the relay of ironic effect in the target text. The subtitler superseded a low-frequency word "contemplative" (BT: "kontemplacyjny") with a more colloquial high-frequency word "Zastanawiania się?" (BT: "Thinking about?") which presumably should be easier for the participants to process.

Another example of modification that has been identified in the descriptive data set is **SH2_1**. In this scene, Sherlock is again being sarcastic and replies to Watson with a personal insult to mock him with the utterance: "Wolałbyś taki nietakt, jak porządny strój wojskowego ze szpetnym szalikiem, dziełem rąk twojej narzeczonej?" (BT: "Would you prefer such a gaffe, like a decent outfit of a military man with hideous scarf, handmade by your fiancée?"). In this instance of multimodal irony, the original dialogue is condensed and paraphrased in the subtitles significantly. Since the original dialogue is quite long, the reduction was necessary as otherwise, the target audience could find it difficult to read the subtitles due to fast presentation rate. The subtitler also replaced some individual phrases that seem to be less known or less frequently used in Polish speech. For instance, the L3 expression "fashion faux pas" is simply rendered as "gaffe" (BT: "nietakt") to facilitate the comprehension process most likely. Despite the modifications, the subtitler retained Sherlock's mocking reference to the Watson's military dress code and the heinous scarf made by Mary, which are key to grasp the intended ironic meaning. Thus, it is assumed that the ironic effect is maintained in the target text.

Finally, an example of modification has also been found in **SH2_3**. In this scene, Sherlock's ironic metaphor intended to mock Watson's failure in the original dialogue is not reduced but paraphrased in the target text as "Mówiłem celnie! Ile szans mam ci dawać?" (BT: "Accurately. I said accurately! How many chances do I have to give you?") to make it sound more natural for Polish audience. Here, Sherlock is using the word "windows" to refer to "a window of opportunity" which is the ideal time to act. This idiomatic expression is

conveyed in a rather colloquial way as "a chance" in the subtitles. Despite the substitution of these two words and the omission of ironic metaphor from the subtitles, it seems that the ironic content still conveys the same meaning in the subtitled version, as it is also the case in the original version.

5.4 Relay of Multimodal Irony in the Voiced-Over Texts

As shown above, the source dialogue was modified in the subtitled clips of SH1 and SH2 on several occasions. In this section, the multimodal analysis of descriptive data reveals how multimodal irony as communicated by the filmmakers in the original dialogue is rendered in the voiced-over versions of the two Sherlock Holmes films. As it has also been found to be the case with the subtitles, the instances of irony relay in the voiced-over version of SH1 and SH2 can also be classified into two broad categories, namely, *preservation* (i.e., the wording and the form of the source language dialogue have been preserved in the target language voice-over while rendering ironic utterances) and *modification* (i.e., the wording and the form of the source language dialogue have been modified in the target language voice-over while rendering ironic utterances). The descriptive analysis showed that, in contrast to subtitles, one half of the source dialogues (3 out of 6 clips) is preserved and the remaining half (3 out of 6 clips) is modified in the voiced-over Polish version of SH1 and SH2. Likewise, preservation in the voiced-over version does not involve any degree of manipulation of the source texts and each word of the original dialogue is transferred. In the following subsection, it has been debated whether the ironic effect that the filmmakers communicated through the verbal modes has been retained where the wording of the source language dialogues has been preserved in the target language s voiced-over clips.

5.4.1 Preservation of multimodal irony in the voiced-over texts

Similarly, to the subtitled versions, preservation is achieved in the voice-over through the transference of each word of the original dialogue. The voice-over translation is known to be "a trustful transfer mode" reflecting "a faithful, literal, authentic and complete version of the original audio" (Franco et al., 2010: 26). A stronger tendency towards verbatim translation is also visible in the analysis below. In this subsection, three instances of multimodal irony have been exemplified and analysed that have been preserved on a lexical level from the SL into the TL in the clips under study.

Clip **SH1_2** is the first illustration of ironic utterance relayed in the voiced-over version with the same set of expressions used in the original dialogue. In subtitles and voice-

over alike, Watson's ironic riposte is rendered into Polish as "Wyglądasz cudownie" (BT: "You look gorgeous") following closely the wording of the source language. It is in accord with a general practice that the translator should strive to achieve an exact correspondence between the target and source texts as much as possible, and particularly when the original soundtrack begins and ends. Although synchrony is not particularly required in the voice-over translation, in this instance, voice-over isochrony helps to "enhance comprehension and maintain characters' defining traits" (Sepielak and Matamala, 2014: 154). The preservation of Watson's ironic riposte thus helps to retain the manifestation of his attitude of dissociation towards Sherlock, as it was originally intended by the filmmakers in the original dialogue.

Similarly, in clip **SH2_2** each word of the original dialogue is also transferred in the voiced-over excerpt. That is to say, Sherlock's mocking dialogue directed to Lord Cowards is conveyed into Polish as "A więc nie ma czasu do stracenia, prawda" (BT: "So there is no time to lose, right?). Interestingly enough, in this scene, the question tag is maintained in the voiced-over excerpt. In the instance of the modified subtitles, it has been highlighted that question tags are not used in Polish speech as often as in English, and hence it has been deleted (and also due to some patio-temporal constraints) most likely. The reason why the question tag is preserved in the voice-over translation can be two-fold. First, the preservation in voice-over is intended to maintain the illusion of authenticity and reality of the original soundtrack (Franco et al., 2010). The transferrence of the question tag which is more common in the English language may thus help to maintain this illusion of reality created in the source text, which is achieved through synchrony, in other words, through the translation of audible phrases (Franco et al., 2010). Second, the spoken language typically conveys more information than the written language. As a result, the translator probably estimated that the preservation of the question tag would not impede the comprehension process of the scene significantly, as the *lektor* will still stop narrating before the original dialogue is finished and the intended ironic meaning will be still intelligibly relayed.

Consider also **SH1_3** where the original dialogue was preserved in the voice-over version. That is to say, Mycroft's dissociative attitude is conveyed into the voice-over as "Szkoda, że przyjaciele nie przyszli" (BT: "Shame your friends did not come") to mock Watson's stag night. In this instance, the translation is in accord with kinetic synchrony and action synchrony (Orero, 2006). The term kinetic synchrony means that "the message read by the voice which delivers the translation must match the body movements which appear on screen" (Orero, 2006: 257). In this case, the voice-over narration correlates with Mycroft's head movement, gestures and facial expression. As far as action synchrony is

concerned, the action and the narrator's voice on the screen have to match. This is also the case in SH1_3 where the voice-over dialogue starts right after Mycroft comes over to Watson's table and finishes before the on-screen characters burst out laughing. A longer or shorter duration of the voice-over recording might distract the viewers from spotting the presence of "antiphonal laughter" (Bryant, 2011) which is highly important in this scene. Thus, it appears that the preservation of each word of the source dialoguehelped to relay ironic meaning in this clip that the film directors aimed to communicate. Sometimes, however, the modification of the source text is indispensable in voice-over, as illustrated in the excerpts below.

5.4.2 Modification of multimodal irony in the voiced-over texts

As presented in Section 5.4, , the remaining half of the source utterances is modified in some way in the voiced-over versions of SH1 and SH2. Here, modification also involves various transformations of the SL text into the TL text resulting in the paraphrased, condensed or reduced voiced-over translations at a textual or sentence level. According to Hołobut (2016: 249) utterances in the voice-over script should remain "conversational and concise, colloquial and polished, flat and singular yet be representative of the polyphony of well-rounded characters", and these features require a number of alterations in the TL especially when the original dialogue involves expressions that the target viewers may be not fully familiar with. In the following examples, it has been outlined what modifications have been made and it has been debated whether the changes made to the source dialogue influence the transfer of ironic meaning in the voice-over.

SH1_1 is an example of modification in which the source utterances are in the voiced-over clip. In this scene, Sherlock's sarcastic comment is rendered into Polish as: "Na jakim etapie? Zastanawiania się?" (BT: "At what stage? Thinking about?"). As it has been found to be the case in the subtitled clip, the phrase "process" is eliminated in the voice-over, while a long phrase "contemplative" is paraphrased into a more informal and shorter equivalent expression "thinking about" to enable the audibility of the original soundtrack. As a result of reduction and paraphrase, the viewers still have a chance to grasp Sherlock's intonation and tone of voice which is of significant narrative importance in this fragment.

As it has been the case in the subtitled version, the original dialogue in clip **SH2_1** has also been modified in its voiced-over counterpart. In this sequence, Sherlock's sarcastic comment is translated into TL as "Wolisz schludny strój wojskowego w połączeniu ze szpetnym szalikiem wydzierganym przez narzeczoną?" (BT: "Do you prefer a neat military

outfit in combination with an ugly scarf knitted by your fiancée?"). Similarly, to the subtitled version, the audiovisual translator also opted for the significant condensation of the original utterance. Although the economy of space and time is not as rigid in voice-over as in subtitles, the voice-over translation is also known as a constrained form of translation that requires a high degree of condensation or reduction of the source text to increase access to the source dialogue (Woźniak, 2012). Since the voice-over *lektor* starts the narration a few syllables after the source text and ends the narration before the on-screen character finishes the dialogue, the timeframe for the voice-over lines is shortened (Schjoldager, 2008). This gives rise to a possible reason why the translator superseded longer and less frequently used expressions in L3, i.e., "fashion faux pas" were replaced with a more conversational, simpler and shorter word "gaffe" (BT: "nietakt"). Since the target translation still relays the references about Watson's military dress code and handmade scarf by Mary, it is assumed that the ironic meaning remained intact, as it was communicated in the original dialogue. In long dialogues, the script is often reduced even by half to enable the *lektor* convey the necessary information (Woźniak, 2012).

A similar example of modification that has been retrieved from the descriptive data set is **SH2_3**. Following the convention to decrease the linguistic complexity and to use more conversational and everyday language instead of idiomatic expressions, the translator also simplified the vocabulary used in the original dialogue with: "Mówiłem celnie! Ile szans mam ci dawać?" (BT: "Accurately. I said accurately! How many chances do I have to give?"). As it was also the case in the subtitled version, the translator also rendered the SL word "windows" into more colloquial expression "chances" in the TL. Thus, Sherlock's ironic metaphor communicated in the original dialogue is again no longer present in the target text, as "chances" have here a very literal meaning. Nevertheless, it is still very likely that the ironic effect is retained in the voice-over, which certainly would not be the case if the original dialog was literally preserved in the target text.

5.5 Multimodal Irony in the Subtitled and Voiced-over Texts

As demonstrated in the sections above, film dialogue cannot serve as the only source of information in the construal of multimodal irony. It is thus essential to complement the verbal modes with the non-verbal acoustic, visual and kinesic modes to compose the meaningful whole. Particularly, when the source text is reduced or condensed in the subtitled or voiced-over clip, the viewers may need to resort to the information relayed visually and acoustically to enhance their understanding of the intended ironic meaning through the non-

verbal semiotic resources. However, the perception of these non-verbal visual cues can be determined by the level of proficiency in the source language. In this section, the relationship between the modality of audiovisual translation and levels of English proficiency in the comprehension process of multimodal irony in the excerpts under study is considered. It is discussed how the participants with high (HLPs), medium (MLPs) and low (LLPs) level of English language competence are expected to comprehend multimodal irony relayed with subtitles and voice-over in the two Sherlock Holmes films.

5.5.1 Clip: Cemetery and Lord Blackwood

In addition to the verbal modes, clip **SH1_1** features the use of the acoustic, kinesic and visual modes in the generation of ironic meaning. In this scene, the acoustic modes, and particularly voice modulation, are vital for the expression of Sherlock's dissociative attitude. It is also supported by the kinesic and visual modes including facial expressions, head and body movements as well as camera distance and angle. Multimodal irony is modified both in its subtitled and voiced-over versions in TL. As the source dialogue was partially reduced and paraphrased in the subtitles, I would expect HLPs and MLPs to skip the subtitles significantly and predominantly focus on the Sherlock's face to retrieve his tone of voice as well as features of facial expression. Due to a low level of English, it would be assumed that LLPs will fixate on the subtitles for longer, and as a result will not spot the non-verbal visual and kinetic semiotic resources considerably. In the voiced-over version, I would anticipate that all participants remained focused on the on-screen characters' faces and the upper part of their bodies in medium-close-up shots to extract more information. These gaze patterns will be tested in the next chapter.

5.5.2 Clip: Watson and Sherlock

In **SH1_2,** the presence of the kinesic and visual modes is particularly important, while the acoustic modes play a rather supportive role in the generation of ironic meaning multimodally. In this clip, Watson's attitude of dissociation is predominantly construed visually through mise-en-scène (settings, costumes, Watson's body language and facial expressions), cinematography (long and medium close shot), editing (cut-in technique) and, to a smaller extent, through Watson's voice quality. Watson's original dialogue was preserved both in the subtitled and voiced-over versions of SH1. The experimental analysis is expected to show how participants split their visual attention across the subtitled and voiced-over screen. Since the subtitled text is rather short, I would expect that HLPs, MLPs,

136

and LLPs will avoid looking at subtitles and will principally dwell on Watson's and Sherlock's face and body. In the voiced-over clip, I would predict to see even higher fixation values on the faces and bodies of the on-screen characters due to the absence of subtitles. These assumptions will be verified with the eye-tracking technology in the next chapter.

5.5.3 Clip: Sherlock and Lord Coward

In **SH1_3**, multimodal irony is composed by the acoustic, kinesic and visual modes and particularly acoustic modes are instrumental as they highly contribute to the manifestation of Sherlock's dissociative attitude to the proposition echoed through the music score in combination with Sherlock's flat intonation and slow speech rate. The visual and kinesic modes play here a more ancillary role construing irony mainly through Sherlock's facial expressions such as a blank face, raised eyebrows or wide-opened eyes in medium close-up shots. The source dialogue was preserved in the voice-over version but modified in its subtitled counterpart through the omission of the question tag. Although the information conveyed acoustically is highly significant in this clip, I would anticipate that HLPs, MLPs and LLPs viewers avoid looking at the subtitles and predominantly gaze at Sherlock's face to grasp the peculiarities of his facial expressions rather easily. Similar viewing behaviour I would assume to see in HLPs, MLPs and LLPs watching the voiced-over clip to recognise Sherlock's mocking attitude towards the proposition echoed by Lord Coward.

5.5.4 Clip: Sherlock, Watson and a scarf

Clip **SH2_1** gives prominence to the visual and kinesic modes, while the acoustic modes are considered as supportive in the production of the intended ironic meaning. In this scene, Sherlock's attitude of dissociation is predominantly expressed through a combination of Sherlock's body language, costumes, props, eye contact, camera position and angle as well as cut in and out technique, while Sherlock's intonation, volume, pitch are of secondary importance. Sherlock's original dialogue is here reduced and considerably rephrased both in the subtitled and voice-over version of the Sherlock Holmes film. Due to the length of the subtitles and a number of changes made to the content of the subtitled text, I would assume that some of HLPs and the majority of MLPs and LLPs will spend the vast amount of time at the bottom of the screen processing the content of the modified subtitles. Consequently, they will miss the non-verbal and kinetic semiotic resources in the middle of the screen to a large extent. In the voiced-over clip, on the other hand, I would expect that all participants

will take advantage of the absence of subtitles and will split their visual attention between Sherlock's face and body in order to interpret the intended meaning as ironic.

5.5.5 Clip: Watson's stag party

Clip **SH2_2** also features the use of the visual, kinesic and acoustic modes along the verbal modes in irony relay. In this scene, the acoustic modes are particularly important as both diegetic and non-diegetic sounds are instrumental to the retrieval of irony with the help of Mycroft's voice quality and "antiphonal laughter" (Bryant, 2011) of all the on-screen characters. The information relayed visually such as Mycroft's facial expressions, head and body movements or camera angle are recognised as supportive in the expression of Mycroft's dissociative attitude. Each word of the original dialogue was preserved in its subtitled and voiced-over versions in TL. In the voice-over clip, I would anticipate that all HLPs and MLPs would largely avoid gazing at the subtitles and will largely focus on Mycroft's face with intention to interpret his voice quality and "antiphonal laughter". LLPs will probably find it more challenging to divide their attention between reading the subtitles and following the action in the middle of the screen, and as a consequence will not spot peculiarities of Mycroft's facial features. The perception of the non-verbal visual and kinesic semiotic resources should be more straightforward in its voiced-over counterpart. Their viewing behaviour will be compared and contrasted in the following chapter.

5.5.6 Clip: The ambush on a train

In clip **SH2_3**, multimodal irony was principally construed by the visual and kinesic modes and to a smaller extent by the acoustic modes. In this scene, Sherlock's attitude of disapproval is predominantly construed visually through a collaboration of fast and slow sequences, abrupt and dramatic shots, dynamic camera movements and body language, while Sherlock's voice modulation is considered ancillary. The original dialogue is paraphrased both in the subtitled and voiced-over clip. Since the idiomatic expression in the source text is superseded with a more colloquial phrase in the target text, I would expect that a great number of HLPs and MLPs, unlike LLPs, will avoid looking at subtitles and will retrieve ironic meaning through the perception of various non-verbal semiotic resources. In the voiced-over version, I would predict that all viewers will principally dwell on Sherlock's face with the attempt to spot his emotions and facial expression as well as his tone of voice, although the narrator hinders the reception of the original dialogue substantially in this scene. I will confirm my predictions with the experimental component in the next chapter.

5.6 Conclusions

The main goal of Chapter 5 was to examine the way in which the verbal and non-verbal modes are combined to produce meaning on screen, and specifically ironic meaning. Additionally, it aimed to gauge to what extent these verbal and non-verbal modes of film language contribute to irony relay in the selected subtitled and voiced-over clips of *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011). Overall, the multimodal analysis of the scenes featuring the use of multimodal irony strengthens the idea that irony in the film text is a multimodal phenomenon construed by the synergy of the verbal and non-verbal modes.

As presented in Section 5.3, the verbal mode plays a highly significant role in the construal of irony on the basis of the examples considered in this chapter. However, it cannot construe or convey ironic meaning on its own if isolated from the visual, kinesic and acoustic modes. As a result, the same sentence may lose its ironic meaning when expressed without any accompanying non-verbal resources. The linguistic cues and other verbal manifestations like repartee, humour or riposte thus "require the supportive presence of other contextual sources, especially marked intonational patterns, since many of these constructions may also be used in non-ironic contexts besides their *possible* ironic use" (Yus, 2002: 11, emphasis as in original).

As shown in this chapter, the verbal in tandem with the non-verbal semiotic resources of the film language like gestures, clothes, settings camera movements or editing techniques have to jointly contribute to the expression of dissociative attitude, and thus composition of multimodal irony successfully on screen. The evidence from this analysis corroborates Pelsmaekers and Van Besien's (2002) initial claim that the viewers who process verbal and non-verbal sem iotic resources simultaneously may conclude that speakers' utterances should not always be taken literally. Although relay of the ironic intent into the TL is significantly reduced or paraphrased in several instances of the subtitled and voiced-over versions, as the analysis in subsections 5.4 and 5.5 indicates, subjects in this study still have a range of possible non-verbal resources at their disposal such as music, kinesics, camera angle or editing that can compensate for the semantic and linguistic loss of the original dialogue in the TL versions.

Nevertheless, both subtitling and voice-over can hinder the audience perception of certain non-verbal resources that may be salient in the irony comprehension process. That is to say, subtitles may drag the viewers' attention from the action in the middle of the screen when they keep looking at the subtitles. As a result, the audience may fail to grasp a smirk,

gestures, body movements, props, camera angle or editing shots that enhance the interpretation process. Voice-over, on the other hand, can cover the film score or the on-screen character's voice modulation, both of which often play a pivotal role in the construal and relay of multimodal irony, as shown in Section 5.2. The *lektor*'s voice overlapping the original dialogue may therefore hinder the viewer's understanding of ironic meaning.

As mentioned in Section 5.4, comprehension of multimodal irony in the subtitled and voiced-over version of SH1 and SH2 may also be determined by viewers' proficiency level in English in the context of this study.

In the next chapter, the results obtained from the eye-tracking experimental are outlined in order to verify the data obtained from the multimodal transcription analysis presented in this chapter. Chapter 6 thus reveals how HLPs, MLPs and LLPs allocated their visual attention on screen when watching the subtitled and voiced-over clips of SH1 and SH2 in regard to the modality of translation and the level of English proficiency.

6 EXPERIMENTAL COMPONENT: EYE TRACKING

6.1 Introduction

The previous chapter discussed how multimodal irony is construed in the selected subtitled and voiced-over excerpts of SH1 and SH2 from a multimodal analytical approach. As a result of the analysis, a numerous verbal and non-verbal resources were distinguished pertaining to the language of film. In Section 5.5, I have debated whether and/or to what extent viewers participating in the experimental component of this thesis would be able to discern these non-verbal cues in the middle of the screen when watching the subtitled and voiced-over clips. I have also made some predictions about the gaze patterns of HLPs, MLPs, and LLPs in relation to their abilities to allocate their attention between the area of interest around the subtitles and/or the faces and bodies of the on-screen characters. Nevertheless, the results of this multimodal analysis were the fruit of theoretical knowledge and remain to be tested with data from participants. In this thesis, a multi-method approach has been applied to record all the viewers' eye movements, and concurrently to collect their responses using a questionnaire as a methodological tool.

This chapter presents the principal findings of the eye-tracking data of viewers who watched the subtitled and voiced-over clips included in the data set. This includes an overview of their viewing behaviour, comparing the ways they distributed their visual attention on screen to retrieve the intended meaning according to their level of English proficiency.

To date, a number of eye-tracking studies have been conducted examining the visual behavior of participants with different levels of language skills while watching subtitled captions. In one of the first experiments, d'Ydewalle et al. (1987) did not distinguish much variety in gaze patterns among viewers representing different linguistic abilities and reported that "the subjects who know very well the spoken language still read the subtitles as much as the other subjects" (1987:320). This view has been challenged by a more recent study conducted by Orrego Carmona (2015) who showed that there are clear differences between participants with high and low level of proficiency in the source language in terms of attention allocation and these two groups exhibited different types of visual behaviours. That is to say, the viewers with high level of SL fixated less on the subtitle area, while the viewers with low level of SL looked at the subtitles much more frequently. Following these lines of reasoning, I will examine whether the participants of my experimental component behave

differently depending on their level of proficiency in English, and thereby I will test the following hypotheses: (i) HLPs and MLPs made shorter fixations and LLPs made longer fixations in the area of interest around subtitles; (ii) HLPs and MLPs made longer fixations and LLPs made shorter fixations in the area of interest around faces of the on-screen characters in the subtitled excerpts. In the voiced-over counterparts, I will analyse whether the participants exhibit more homogenous visual patterns irrespective of their level of English, and thereby I will verify the following predictions: HLPs, MLPs and LLPs fixate in the AOI around faces of the on-screen characters in similar proportions. This gaze pattern finds its grounding in previous studies (e.g., Hershel and Hochstein, 2005; Treuting, 2006) highlighting that face is the part of the screen that attracts the greatest attention in the middle of the screen in the non-subtitled conditions. In addition, it has also been assumed that HLPs, MLPs and LLPs spend more time gazing at the areas of interest around the bodies of the on-screen characters in the absence of subtitles.

The chapter consists of three main sections. Section 6.2, examines the ways in which the target audience watched three subtitled and voiced-over clips in which multimodal irony was principally construed visually, while Section 6.3 looks at the visual behaviour of participants in three other subtitled and voiced-over clips featuring multimodal irony that is construed mainly acoustically. This chapter finishes with Section 6.4 in which the most significant findings derived from this eye-tracking study are discussed in light of the Polish audiovisual scene.

6.2 Visually Construed Multimodal Irony

As explained in Section 5.2, there is a wide array of non-verbal visual resources that are considered essential in the construal of multimodal irony in the clips under scrutiny. Among others, a variety of facial expressions, gestures, body movements, garments, props as well as camera motifs and editing techniques were identified as contributing to the relaying of an attitude of dissociation on screen.

This section discusses how Polish participants with three different levels of English proficiency allocated their visual attention across several areas of interest (AOIs) which each contain non-verbal information primarily relayed visually with supporting information conveyed acoustically. In addition, it compares how theseviewers switch their focus of attention between the AOI around the subtitles and the AOIs in the centre of the screen to grasp the intended meaning on the screen.

6.2.1 Analysis of clip: Watson and Sherlock

In SH1_2, as expected, all participants had a chance to grasp a whole range of non-verbal resources pertaining to mise-en-scène to support their interpretation of Watson's dissociative attitude towards Sherlock while subtitles were displayed at the bottom of the screen. The eye-tracking data in Tables 6.1 and 6.2 show that, in the subtitled version, Sherlock's body and Watson's face maintained visual attention of a great majority of participants in the middle of the screen. Unsurprisingly, LLPs fixated on the subtitles more than other subjects, as shown in Tables 6.1 and 6.2.

Table 6.1 Number of fixations, total number of fixations with the percent ratios in the AOIs in SH1_2 depending on the level of English proficiency (high, medium and low) in the SUB_VO group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	22	40	30	43	17	31
Sherlock's face	8	15	10	14	12	22
Subtitles	3	5	8	11	13	24
Watson's body	3	5	5	7	3	5
Watson's face	19	35	17	24	10	18
Sum	55	100%	70	100%	55	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	5888	40	6368	40	3263	27
Sherlock's face	1980	14	2573	16	3553	29
Subtitles	524	4	1206	8	2227	18
Watson's body	591	4	849	5	499	4
Watson's face	5680	38	4930	31	2554	22
Sum	14663	100%	15926	100%	12096	100%

Table 6.2 Mean fixation duration in the AOIs in SH1_2 depending on the level of English proficiency (high, medium and low) in the SUB_VO group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Subtitles	174.67	150.75	175.15
Sherlock's face	236.25	257.30	296.08
Sherlock's body	267.64	212.27	191.94
Watson's face	298.95	169.80	255.40
Watson's body	197.00	290.00	166.33

As shown in Table 6.1, HLPs and MLPs demonstrated a number of common gaze patterns. For instance, both HLPs and MLPs largely avoided looking at the subtitles. That is, only 5% and 11% of their fixations were made on the subtitles, which they read for slightly more than

500 and 1200 ms, respectively. These low fixation values suggest that HLPs and MLPs process them under minimal cognitive load. This visual behaviour suggests that the viewers with high and medium level of English proficiency do not have to rely on the written mode entirely to grasp the intended meaning in this excerpt as they can largely retrieve it from the original soundtrack in the source language. Hence, they looked at the subtitled text rather sporadically probably as a form of re-confirmation of information they misheard or failed to understand, as it has also been the case in previous studies (e.g., Orrego-Carmona, 2015). Another possible explanation for the short fixation duration in the subtitle area of interest can be linked to the fact that the participants' eye movements are influenced by their reactions to bottom-up stimuli, and hence look at the subtitles automatically as they appear on screen without activating their cognitive processes (Webb & Renshaw, 2008: 39). Interestingly, however, on average HLPs read the captions longer than MLPs, that is 174 and 150 ms, respectively, as the mean fixation durations in Figure 6.2 illustrates. LLPs, on the other hand, needed more fixations (25%) and more time (2227 ms)respectively, and thus invested more mental effort than HLPs and MLPs in reading the subtitled text, thus providing evidence for increased cognitive load.

While low level of foreign language skills would appear to account for the increased fixations values, let us consider several other factors that could influence this type of viewing behaviour. The original dialogue in SH1_2 was preserved in the Polish one-line subtitles which are known to attract fewer and shorter fixations. Although this type of subtitles should be therefore easier to process (Ghia, 2012), some of the viewers, in particular LLPs, spent more time reading them than expected. The main reason for this is that some words within the subtitled text require more or less processing time. For example, high-frequency words are easier to process than low-frequency ones (Kruger, Szarkowska and Krejtz, 2015). Although the word "cudownie" in thesubtitles (BT: "gorgeous") can be assigned to the group of high-frequency words, it appears that the participants dwelled on it the most, as the heat map visualisations show in Figure 6.1 below.

Figure 6.1 Heat maps for HLPs, MLPs and LLPs in SH1_2 at 00:01:14
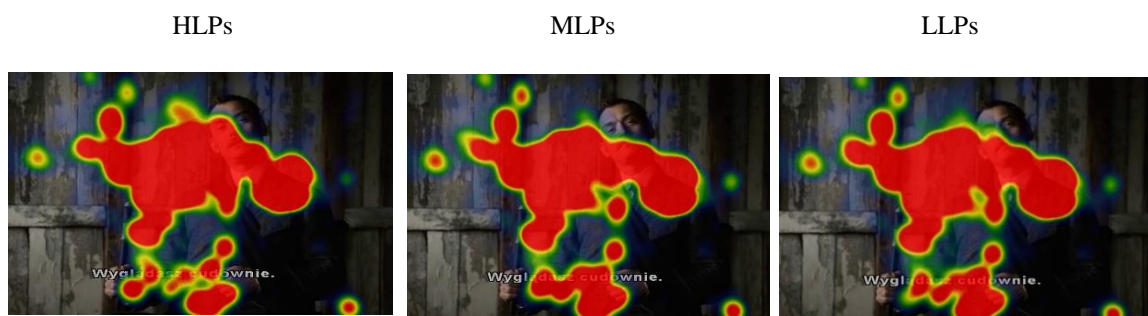
| HLPs | MLPs | LLPs |

The heat maps in Figure 6.1 illustrate the difference in the duration and number of fixations resulting from the data of HLPs, MLPs and LLPs. As can be seen, in the heat maps above, higher concentration of fixations is visible in the beginning of the word "cudownie" within the subtitle line. The word "cudownie" is here of particular importance to the narrative. It can be classified as an emotion word with a positive affective valence. In this case, however, the word "gorgeous" is the opposite of what Watson actually intends to say and the understanding of the filmmakers' intentions requires from some participants a higher level of cognitive load to process the intended meaning. LLPs have to juxtapose Watson's dialogue with Sherlock's appearance to fully comprehend the ironic meaning in this scene. Otherwise, they run a risk of grasping the literal or positive meaning of Watson's ironic riposte (see Nilsson, 2013; Vandaele, 1999/2002), which could lead to a false interpretation of his dissociative attitude against Sherlock. Thus, the quality rather than quantity of words used in the subtitled text would seem to make a notable difference in the fixation distribution in the subtitle area, as it has also been shown in other studies (e.g., Dwyer, 2015). The processing of emotional expressions like "cudownie" may also be more challenging for LLPs due to their low level of English proficiency, as they are less familiar with the foreign language and culture. More empirical research on the effect of language proficiency on the perception of emotion words is required to confirm this insight.

The multimodal irony relayed via subtitles would certainly be misunderstood, if the verbal modes were not supported by the co-existence with several non-verbal elements. As the one-line subtitles do not provide much "more information than what can already be extracted from the image and the auditory message" (d'Ydewalle and de Bruycker, 2007: 202), all participants, including LLPs to a great extent, managed to shift their attention away from the verbal resources (subtitles) towards the kinesic (e.g., facial expressions) and visual resources (e.g., props). That is to say, HLPs, MLPs and LLPs found the AOI around Sherlock's body visually the most attractive, as the percent ratio of fixations (40%, 43% and 31%, respectively), and the total fixation durations (5888, 6368 and 3263 ms, respectively) clearly indicate. It means that Sherlock's body position and Sherlock's clothes contained visual information highly important to the narrative content of the scene, and viewers had an opportunity to spot the details that supported their comprehension of Watson's attitude. The AOI around Watson's face also attracted a lot of attention from HLPs, MLPs and LLPs, as eye-tracking data in Tables 6.1 and 6.2 indicate. These gaze patterns enabled all participants to perceive some peculiarities of Watson's facial expressions (e.g., smirking) that were identified as essential to the comprehension of Watson's dissociative attitude in the

descriptive strand of the analysis. Generally, the total fixation durations in the subtitle area is still rather small in comparison to the fixation values in the middle of the screen. The target audience, including LLPs, thus spent a significant amount of time exploring the AOIs in the centre of the screen, despite the fact that subtitles were displayed in the bottom. Concurrently, participants had access to the original soundtrack (e.g., Watson's voice modulation) that could additionally contribute to the comprehension of the intended meaning. In the voiced-over version, on the other hand, Sherlock's body and Watson's face still remained of great interest for all participants, as indicated by the number of fixations as well as the total fixation fixations durations in Tables 6.3 and 6.4. Whether the voice-over narration posed challenges to retrieve Watson's voice quality will be discussed in the next chapter.

Table 6.3 Number of fixations, total fixation duration with the percent ratios in the AOIs in SH1_2 depending on the level of English proficiency (high, medium, low) in the VO_SUB group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	27	51	21	36	23	43
Sherlock's face	9	17	10	17	7	13
Watson's body	2	4	1	2	1	2
Watson's face	15	28	26	45	22	42
Sum	53	100%	58	100%	53	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	7279	42	5805	28	6597	38
Sherlock's face	2758	16	4700	23	3142	18
Watson's body	666	4	216	1	133	1
Watson's face	6769	39	10010	48	7498	43
Sum	17472	100%	20731	100%	17370	100%

Table 6.4 Mean fixation duration in the AOIs in SH1_2 depending on the level of English proficiency (high, medium and low) in the VO_SUB group

| Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Sherlock's face	259.96	335.71	392.75
Sherlock's body	306.44	263.86	263.88
Watson's face	451.27	348.28	340.82
Watson's body	333.00	108.00	133.00

As can be seen from Table 6.3, similarly to the subtitled version, Sherlock's body and Watson's face also drew and captivated the greatest attention from all participants when watching the clip with voice-over underlying the narrative importance of the non-verbal resources enclosed in these AOIs. Since the values presented in the form of number and total duration of fixations in the AOIs are higher in the centre than in the bottom of the screen, the target audience could supposedly identify more non-verbal elements of film language used to construe Watson's attitude. The identification of these non-verbal cues could thus help them recognise the emotional engagement between the on-screen characters more successfully in the voiced-over than in the subtitled version of the clips as they were not distracted by the presence of subtitles on the screen. It can be also seen from the data in Table 6.3 that Watson's face attracted many more and longer fixations than Watson's body in HLPs, MLPs and LLPs confirming results obtained from the previous research that face is the part of the screen that attracts the greatest attention (e.g., Treuting, 2006). These high fixation values as well as the mean fixation durations in Figure 6.4 indicate that the subjects attempted to grasp characteristic traits and features of facial expressions to recognise Watson's emotions and intentions, and thus his attitude of disapproval towards Sherlock (Herdández et al., 2009). In regard to the AOI around Sherlock, interestingly, HLPs, MLPs and LLPs spent much less time looking at Sherlock's body than his face, as Figure 6.3 shows. As a result, it appears that the participants in this group found the visual narrative cues embedded in the AOI around Sherlock's body (e.g., Sherlock's posture and clothes) more helpful to spot the incongruity between Watson's dialogue and Sherlock's looks to come to the conclusion that Watson's statement is not expected to be understood as literal. Thus, whether the participants focus more on face or body of the on-screen characters is predominantly determined by their importance to the narrative. Nevertheless, these fixation values on the on-screen characters' faces and bodies in the voiced-over version were similar to the ones in its subtitled counterpart, despite the display of subtitles on the screen. It is therefore likely that some of the viewers, particularly HLPs and MLPs, will not extract more information conveyed visually from the voiced-over than from subtitled excerpt, as illustrated in the next chapter, which could lead to increased comprehension scores. In addition, the voice-over narrator covered the majority of the original soundtrack drowning the supportive function of acoustic cues like Watson's intonation and pitch, which may impact the irony comprehension scores to a greater extent than the appearance of subtitles in the bottom of the screen.

6.2.2 Analysis of clip: Sherlock, Watson and a scarf

In **SH2_1**, as anticipated, some of the participants experienced challenges to follow the action in the middle of the screen in the subtitled version. As the eye-tracking data in Tales 6.5 and 6.6 revealed, major discrepancies were found in viewing behaviour among HLPs, MLPs and LLPs. In this instance of multimodal irony, subtitles highly competed with other AOIs in the centre of the screen for visual attention. The subtitled text dragged particularly MLPs' and LLPs' attention from the perception of nonverbal cues conveyed visually on the screen. As a result, the participants could fail to spot body movement or elements of garments that are expected to support the detection of Sherlock's dissociative attitude to the proposition echoed, mocking Watson.

Table 6.5 Number of fixations, total fixation duration with the percent ratios in the AOIs in SH2_1 depending on the level of English proficiency (high, medium, low) in the VO_SUB group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	**HLPs**		**MLPs**		**LLPs**	
Sherlock's body	11	6	12	5	14	6
Sherlock's face	64	37	49	22	39	17
Watson's body	2	1	5	2	5	2
Watson's face	51	30	51	23	46	20
Subtitles	43	25	102	47	128	55
Sum	171	100%	219	100%	232	100%
Measurement	**Total fixation duration (ms)**	**Percent ratio**	**Total fixation duration (ms)**	**Percent ratio**	**Total fixation duration (ms)**	**Percent ratio**
AOIs	**HLPs**		**MLPs**		**LLPs**	
Sherlock's body	3022	6	3615	7	4157	7
Sherlock's face	23753	48	15896	30	10118	18
Watson's body	350	1	803	2	998	2
Watson's face	15406	31	14931	28	15515	28
Subtitles	6946	14	17558	33	25592	45
Sum	49477	100%	52803	100%	56380	100%

Table 6.6 Mean fixation duration in the AOIs in SH2_1 depending on the level of English proficiency (high, medium and low) in the VO_SUB group

Measurement	Mean fixation duration		
AOIs	**HLPs**	**MLPs**	**LLPs**
Subtitles	157.86	173.84	199.94
Sherlock's face	371.14	324.41	285.41
Sherlock's body	279.43	0.00	0.00
Watson's face	302.08	292.76	328.59
Watson's body	175.00	160.60	199.60

Table 6.5 illustrates that the split in attention between the subtitle area and the AOIs in the middle of the screen is immediately visible. For example, the total fixation duration shows that HLPs dwelled on the subtitles much longer in clip SH2_1 (nearly 7000 ms) than in SH1_2 (524 ms), which is more likely to be the result of the subtitle length rather than the cognitive effort necessary to process them. HLPs like other fluent readers of subtitles, as revealed in the questionnaire (see Appendix 8), "have been found to have no difficulty following subtitles" (Kruger, Szarkowska and Krejtz, 2015), and thus probably did not utilize much considerable cognitive effort to comprehend irony relayed via subtitles as this information they should be able to retrieve from the original soundtrack, which will be verified in the next chapter. As a result, their eye movements could therefore be guided by the length of the two-subtitle lines rather than by the cognitive load involved in the processing of the subtitled text. On average, HLPs were also dwelling on the subtitles for a shorter period of time that MLPs and in particular LLPs, as anticipated (Table 6.6). These remaining two subgroups of viewers, on the other hand, relied more heavily on the subtitled text. That is to say, MLPs and LLPs had 47% and 55% of their fixations located on the subtitles, respectively, and kept reading them much longer than HLPs, that is, 17558 and 25592 ms, respectively, which equals to 33% and 45% of their total time (Table 6.5) spent on the AOI around the subtitles in comparison to other non-verbal AOIs in this clip. The fact that MLPs and LLPs had about half of their fixation values on the subtitle area suggests that the subtitle length is probably only one of the factors responsible for their total fixation duration, and in addition the ironic meaning conveyed via subtitles could be particularly difficult for these groups of viewers to process. While MLPs and LLPs viewing behaviour is closely related to their greater reliance on subtitles due to their limited English proficiency, there are also other factors that might influence these gaze patterns. In addition to one-line subtitles, the original dialogue in SH2_1 was also rendered into two-line subtitles, which are known to involve more regular and word-for-word reading behaviour. Hence, as observed above, two-line subtitles automatically attracted more fixations and required from viewers more time to read them. Nevertheless, as mentioned earlier in this chapter, cognitive load is more related to the quality rather than the quantity of words in the subtitled text (e.g., Dwyer, 2015). These gaze patterns are thus rather associated with the narrative saliency of these lexical units that grab the audience's attention to particular visual cues within a scene. That is, the original dialogue was transferred into subtitles with a set of more colloquial phrases such as such as "nietakt" (ST: "fashion faux pas") or "schludny strój wojskowego" (ST: "fine military dress") should sound familiar to the target audience. It does not mean,

however, they will make Sherlock's dissociative attitude easier to detect. MLPs and LLPs also have to process the emotional component of word stimuli (e.g., Sutton and Altarriba, 2008) such as the word "heinous", and interpret it concurrently with the visual and acoustic information. The interpretation of emotional vocabulary can be more difficult for these viewers who are not fully involved in the second language culture. Furthermore, "subtitle translation is also likely to influence the perception of the audiovisual product and viewers' general reading patterns" (Ghia, 2012: 175). In SH2_1, the original dialogue was significantly condensed and modified in the subtitles (see Section 5.4) which can also account for higher fixation values in the subtitle area by MLPs and LLPs. Slower reading, and thus longer fixations, as those made by LLPs in the subtitled text, are also often associated with lower abilities in text comprehension (Schmidt-Weigand et al., 2010) – to the extent that they can impede understanding of the intended ironic meaning, which will be crosschecked in the next chapter.

Although the processing of subtitles limited MLPs and LLPs perception of non-verbal cues, the viewers still managed to dwell on the AOIs in the centre of the screen to a certain extent, as Table 6.5 indicates. As shown by the number and total duration of fixations, MLPs and LLPs also had a chance to extract some information from the AOIs around Sherlock's and Watson's faces, and thereby identify emotions and build a sort of emotional engagement with the main on-screen characters. HLPs, instead, took much more time to explore the AOIs around Sherlock's and Watson's faces which could help them "reveal the incipient thought of the characters" (Redmond et al., 2015), as the fixation values in Table 6.5 suggest, despite the appearance of subtitles. Thus, HLPs could examine Sherlock's and Watson's faces in ample details and spot some traits and features of facial expressions that could help them interpret Sherlock's attitude as ironic. These gaze patterns demonstrate that the presence of subtitles is biased towards the on-screen characters' faces, rather than to the peripheries of the screen, as it is also observed in other studies (e.g., Hershel and Hochstein, 2005).

As far as the voiced-over version is concerned, HLPs, MLPs and LLPs showed overall a similar viewing behaviour, predominantly fixating on Sherlock's and Watson's faces, as Table 6.7 shows. Counterintuitively, the vast majority of participants (HLPs and MLPs) largely avoided looking at the AOIs around Sherlock's body with the exception of LLPs, who divided their attention between Sherlock's face and body almost equally.

Table 6.7 Number of fixations and total fixation duration with the percent ratios in the AOIs in SH2_1
depending on the level of English proficiency (high, medium, low) in the SUB_VO group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	28	16	24	15	46	33
Sherlock's face	99	57	83	52	49	35
Watson's body	1	1	9	6	12	9
Watson's face	45	26	45	28	34	24
Sum	173	100%	161	100%	141	100%
Measurement	**Total fixation duration (ms)**	**Percent ratio**	**Total fixation duration (ms)**	**Percent ratio**	**Total fixation duration (ms)**	**Percent ratio**
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	7866	13	6704	11	11338	23
Sherlock's face	36700	59	33487	57	22029	44
Watson's body	141	0	1524	3	3119	6
Watson's face	17208	28	16858	29	13540	27
Sum	61915	100%	58573	100%	50026	100%

Table 6.8 Mean fixation duration in the AOIs in SH2_1 depending on the level of English proficiency
(high, medium and low) in the SUB_VO group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Sherlock's face	370.41	403.46	449.57
Sherlock's body	280.93	279.29	246.39
Watson's face	382.40	374.62	398.24
Watson's body	141.00	169.33	259.92

Table 6.7 presents that HLPs, MLPs, LLPs shifted a part of their cognitive effort from the visual to the auditory channel due to the voice-over narrator. As a result, the Polish participants searched for relevant visual information predominantly in the AOIs around Sherlock's and Watson's faces more likely to "trace the undulating valleys of emotion" expressed on them (Redmond et al., 2015), and support their interpretation of Sherlock's dissociative attitude. In a close-up shot, all participants spent a significant proportion of time exploring especially Sherlock's face, which is expected to facilitate the identification of facial features that should enhance their understanding of Sherlock's intentions to avoid possible linguistic misunderstandings. The fixation values in Table 6.7also confirm that the mean fixation durations, which demonstrate that on average the participants were also dwelling on Sherlock's and Watson's face the longest.

To my surprise, in contrast to many and long fixations in the AOIs around faces of the on-screen characters, the vast majority of the viewers largely avoided looking at the AOIs

around their bodies, although they were exposed to a long shot before the camera moved to a close shot. What is striking about the data in Table 6.7 is that LLPs dwelled more on the on-screen characters' bodies than the two other subgroups. These gaze patterns suggest that LLPs were looking for narrative cues in Sherlock's kinesics and appearance to facilitate their comprehension process. Interestingly, LLPs also seemed to "take a greater area of focus" with some glances outside of the AOIs to spot other elements of mise-en-scène. The screen shots in Figure 6.2 show a higher concentration of fixations of HLPs and MLPs in the middle of the screen, while LLPs spread their attention also in other directions visibly.

Figure 6.2 Heat maps for HLPs, MLPs and LLPs in SH2_1 at 00:01:00

HLPs	MLPs	LLPs

This more explorative viewing behavior, as demonstrated in Figure 6.2, exhibited by LLPs suggests that they attempted to extract information in a more exploratory way which may be linked to LLPs' limited abilities to recognise emotions, due to their low level of English proficiency. As a result, they kept looking for other semantically significant cues across the AOIs and beyond to compensate this gap.

Although the voice-over narrator hampers a great majority of the original soundtrack, and thereby Sherlock's voice qualities are hardly audible, all viewers highly benefited from the absence of subtitles fixating in the middle of the screen much longer and more frequently. Whether these visual patterns also lead to a better comprehension of Sherlock's dissociative attitude and the intended meaning behind it will be verified in the next chapter.

6.2.3 Analysis of clip: The ambush on a train

Contrary to my expectations, in **SH2_3** overall, the eye-tracking data in Tables 6.9 and 6.10revealed some of the major similarities in the attentional distribution among HLPs, MLPs and LLPs across the AOIs that involved more intense competition for viewer's attention among stimuli in the subtitled version. Since a great deal of the participants overall remained focused on the subtitled text, they might miss some of the visually significant information that is highly relevant in the interpretation of Sherlock's dissociative attitude in this clip.

Table 6.9 Number of fixations, total number of fixations with the percent ratios in the AOIs in SH2_3 depending on to the level of English proficiency (high, medium low) in the VO_SUB group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	1	2	2	3	1	2
Watson's face	20	31	16	24	8	12
Sherlock's body	0	0	0	0	0	0
Sherlock's face	15	23	21	32	26	40
Subtitles	29	45	27	41	30	46
Sum	65	100%	66	100%	65	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	492	3	341	2	91	1
Watson's face	4640	30	5307	32	3009	19
Sherlock's body	649	4	1789	11	1298	8
Sherlock's face	5610	36	6147	37	7432	47
Subtitles	4165	27	3129	19	4034	25
Sum	15556	100%	16713	100%	15864	100%

Table 6.10 Mean fixation duration in the AOIs in SH2_3 depending on the level of English proficiency (high, medium and low) in the VO_SUB group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Subtitles	166.00	182.15	177.73
Watson's face	323.00	331.69	376.13
Watson's body	0.00	170.50	91.00
Sherlock's face	374.00	292.71	285.85
Sherlock's body	0.00	0.00	0.00

It is apparent from Table 6.9 that subtitles attracted a substantial amount of attention from HLPs, MLPs and LLPs who made about half of their fixations in the subtitle area, as the percent ratio for the number of fixations (45%, 41% and 46%, respectively) indicates. This viewing behaviour seems to be associated with the subtitle translation of the original dialogue. That is to say, in SH2_3, Sherlock's dialogue involved the use of an idiomatic expression ("windows of opportunity") that is key to the expression of his mocking attitude towards Watson revealing important information to the narrative. In the subtitles, Sherlock's original dialogue was significantly modified and a word ("windows") was replaced with a non-idiomatic phrase ("szanse", BT: "chances"). As visualised in the scan paths in Figure 6.3 the word "szanse" attracted a great deal of visual attention from the participants at all levels of English language proficiency.

Figure 6.3 Heat maps for HLPs, MLPs and LLPs in SH2_3 at 00:04:56

| HLPs | MLPs | LLPs |

Although all viewers gazed at the word "szanse" a great amount of time, as illustrated in Figure 6.3 their visual behaviour can be motivated by different reasons. For example, the eye movements of HLPs and MLPs can be guided by their interest to examine the discrepancies between the source and target texts, as allowed by their advanced skills in the spoken language comprehension. LLPs relied more heavily on subtitles since it was the only verbal channel of communication fully accessible to them due to their low level of English language proficiency. Although HLPs, MLPs, LLPs looked at the subtitles very frequently, their fixations were much shorter, as shown in the total fixation durations (Table 6.9). These gaze patterns suggest that the subtitled text attracted a lot of attention but did not require much cognitive effort to derive the implicit meaning behind Sherlock's attitude of dissociation. Interestingly, however, the total fixation duration in Table 6.9 illustrates that HLPs were looking at the subtitles even longer than LLPs, which is quite surprising, yet on average HLPs' mean fixation durations were much shorter than the average fixation time that LLPs spend on reading the subtitles. The difference in the mean fixation duration also suggests that LLPs needed to activate more cognitive efforts to process the meaning that HLPs did.

It is also worth mentioning that editing patterns in the film text have a huge impact on the way viewers read subtitles, although they are largely unaware of standard editing techniques used in films (see Kruger, Szarkowska and Krejtz, 2015). For instance, in clip SH2_3, editing not only contributes to the construal of multimodal irony but it can also account for the increased cognitive demands due to fast pace cuts "as viewers have to work harder to sustain the illusion of a continuous whole" (Kruger, Szarkowska and Krejtz, 2015).

Nevertheless, subtitles on their own cannot fully relay Sherlock's dissociative attitude. On several occasions, the on-screen characters' faces have been identified as vital sources of attentional and visual information to the narrative (see SH1_2 and SH2_1 this chapter). In SH2_3, despite the display of subtitles, the participants examined the AOIs around Sherlock's and Watson's faces for almost twice as long as the subtitles. As a result,

all viewers, irrespective of their proficiency in English, had the opportunity to discern some peculiar features of facial expressions that would reveal Sherlock's and Watson's emotions, and thereby contribute to their comprehension of the intended meaning.

Let us now consider viewing behaviour in the voiced-over version of SH2_3. On the whole, HLPs, MLPs and LLPs largely dwelled on Sherlock's and Watson's face in the search for narrative cues in the absence of subtitles, as Table 6.11 illustrates. Despite the absence of subtitles, the participants also fixated in the AOIs around Sherlock's and Watson's bodies to a very small extent, as anticipated in a medium close shot.

Table 6.11 Number of fixations and total fixation duration with the percent ratios in the AOIs in SH2_3 depending on the level of English proficiency (high, medium, low) in the SUB_VO group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	19	27	3	6	7	15
Watson's face	14	20	14	26	24	51
Sherlock's body	1	1	1	2	0	0
Sherlock's face	37	52	36	67	16	34
Sum	71	100%	54	100%	47	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	4932	24	541	4	1023	8
Watson's face	4015	20	3365	26	6211	48
Sherlock's body	166	1	342	3	0	0
Sherlock's face	11279	55	8734	67	5787	44
Sum	20392	100%	12982	100%	13021	100%

Table 6.12 Mean fixation duration in the AOIs in SH2_3 depending on the level of English proficiency (high, medium and low) in the SUB_VO group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Watson's face	286.79	240.36	258.79
Watson's body	259.58	180.33	146.14
Sherlock's face	304.84	242.61	361.69
Sherlock's body	166.00	342.00	0.00

As can be seen in Table 6.11, Sherlock's face again received a great deal of visual attention from HLPs and MLPs, as the percent ratio for the number of fixations (52%, and 67%)and for the total fixations duration (55% and 67%), respectively, demonstrate. This viewing behaviour provides evidence that the AOI loaded with visual information is the most engaging or interesting for these participants, as also supported by the mean fixation duration

in Table 6.12 (Sherlock's body attracted attention from only 1 participant who dwelled on this AOI for longer time). LLPs, on the other hand, switched their attention between Sherlock's and Watson's face in similar proportions, as shown in Table 6.11, in the search for additional sources of information. On the whole, these gaze patterns confirm the semiotic significance of the on-screen characters' faces to the construal of meaning, particularly ironic meaning, on the screen.

Although relatively high fixation values on Sherlock's and Watson's faces also allow Polish viewers to identify emotions of the on-screen characters more accurately, and thus Sherlock's attitude of dissociation, this instance of multimodal irony may be more difficult to interpret to due to the fast-paced action and the rhythm of editing involving abrupt shots. In addition to this, sound effects and voice-over narration in concert with these abrupt editing techniques can additionally complicate the retrieval of the on-screen characters' true intensions. In particular, the voice-over narrator's flat, monotonous voice extensively drowns Sherlock's qualities of voice making the interpretation of Sherlock's feelings, and thereby his attitude of dissociation more challenging. This chapter now moves on to analyse the clips in which multimodal irony has been principally construed acoustically to gauge the impact of voice-over narration on the comprehension of the intended meaning.

6.3 Acoustically Construed Multimodal Irony

As noticed in Section 5.2, there are certain sounds such as music score, sound effects and voice modulation of the on-screen characters that were identified as a crucial "aesthetic device […] employed to direct where viewers looked" (Sita et al., 2015) and contributed to the construal of multimodal irony in the excerpts under analysis. The next section will examine whether and/or to what extent viewers representing different levels of English language proficiency drifted their visual attention to the areas of interest (AOIs) that convey the intended ironic meaning mainly acoustically, with visual cues serving as supporting sources of information, as revealed in the eye-tracking data in the following clips.

6.3.1 Analysis of clip: Cemetery and Lord Blackwood

In **SH1_1**, despite the significance of the acoustic modes, particularly MLPs and LLPs relied on the subtitles quite heavily, which is partially unexpected. It appears that the use of subtitles largely depends on the level of English proficiency represented by viewers. All participants predominantly kept looking at the AOIs around faces of the on-screen characters, as predicted. This gaze behaviour is expected to support the interpretation of Sherlock's dissociative attitude towards Lestrade in addition to features of voice modulation.

Table 6.13 Number of fixations, total fixation duration with the percent ratios in the AOIs in SH1_1 depending on the level of English proficiency (high, medium, low) in the SUB_VO group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Subtitles	29	20	54	30	53	39
Sherlock's body	4	1	2	1	4	3
Lestrade's body	3	1	2	2	3	2
Sherlock's face	55	38	76	42	41	30
Lestrade's face	54	37	48	26	34	25
Sum	145	100%	182	100%	135	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Subtitles	4441	15	8777	22	11176	30
Sherlock's body	473	2	233	1	909	2
Lestrade's body	624	2	241	1	543	1
Sherlock's face	11880	40	18887	47	11681	32
Lestrade's face	11927	41	11665	29	12546	34
Sum	29345	100%	39803	100%	36855	100%

Table 6.14 Mean fixation duration in the AOIs in SH1_1 depending on the level of English proficiency (high, medium and low) in the SUB_VO group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Subtitles	153.14	162.54	210.87
Sherlock's face	216.00	248.39	284.90
Sherlock's body	118.25	116.50	227.25
Lestrade's face	232.33	243.02	350.79
Lestrade's body	208.00	120.50	181.00

Table 6.13 illustrates that the number and total duration of fixations increases, when a level of English proficiency decreases, which confirms that "the degree to which viewers will process the subtitles […] will be determined by the extent to which they need the subtitles" (Kruger, Szarkowska and Krejtz, 2015). For example, HLPs skipped the subtitles more frequently than other viewers. They also spent a relatively small amount of time reading the subtitles (441 ms) in contrast to the two remaining subgroups, which means that they did not need to activate more cognitive resources to retrieve the intended meaning. The fixation values of MLPs and LLPs, on the other hand, steadily increased in the subtitle area, as the percent ratios for fixation number (30% and 39%, respectively, in Table 6.13) and the total fixation durations (22% and 30%, respectively, in Table 6.13), respectively, show. Thus, the amount of processing time required to read and understand the subtitles increased, and so

did the mental efforts required to process the intended meaning. These higher eye-tracking values can be linked to MLPs' and LLPs' ability to recognise and integrate information relayed acoustically with subtitles as "the hearing audience matches what they read with what they hear, checking for correspondence of information and interpreting intonation, tenor and other non-verbal elements of speech" (Kruger, Szarkowska and Krejtz, 2015). These viewing patters can also be supported by the mean fixation durations in Table 6.14, which demonstrate that LLPs on average dwelled much longer on subtitles than HLPs or MLPs suggesting that these subtitles were more demanding for LLPs to process.

While many viewers relied on the subtitled text to a greater or lesser extent, all of them managed to examine the AOIs around Sherlock's and Lestrade's faces to a certain degree, confirming that the presence of subtitles attracts the viewers' attention to the areas of central interest like faces instead of distracting them from the visual content, as also shown in the previous studies (e.g., Smith, 2013). From Table 6.13, we can also see that HLPs, MLPs and LLPs spent a great amount of time looking at Sherlock's and Lestrade's faces. In this instance, the viewers' attention may have been drawn by the characteristics of Sherlock's voice qualities that were considered instrumental to the construal of his attitude of dissociation, which will be tested in the next chapter.

Althoughthe participants with low level of English language proficiency needed more time to process the subtitles than the more advanced viewers (i.e., MLPs and HLPs), it does not mean that LLPs achieve poorer scores in irony comprehension, as "there is no evidence that subtitling leads to poorer comprehension" (Smith, 2015). It is also worth mentioning that the presence of subtitles did not limit their access to the AOIs in the middle of the screen to a great extent, which allowed them to perceive a wide array of non-verbal resources such as facial expressions, props or camera functions significant to the narrative that have been identified as supportive information sources in the detection of the intended meaning in this scene.

When it comes to the voiced-over version, according to my expectations, HLPs, MLPs and LLPs remained focused on Sherlock's and Lestrade's face almost exclusively throughout the duration of data collection, as shown in Table 6.15.

Table 6.15 Number of fixations and total fixation duration with the percent ratios in the AOIs in SH1_1 group depending on the level of English proficiency (high, medium, low) in the VO_SUB

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	4	3	1	1	0	0
Lestrade's body	0	0	1	1	0	0
Sherlock's face	88	58	74	53	97	61
Lestrade's face	59	39	64	46	63	39
Sum	151	100%	140	100%	160	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	661	1	300	1	0	0
Lestrade's body	0	0	150	0	0	0
Sherlock's face	27546	62	26012	53	27809	63
Lestrade's face	16549	37	22601	46	16507	37
Sum	44756	100%	49063	100%	44316	100%

Table 6.16 Mean fixation duration in the AOIs in SH1_1 depending on the level of English proficiency (high, medium and low) in the VO_SUB group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Sherlock's face	313.02	309.67	286.69
Sherlock's body	165.25	300.00	0.00
Lestrade's face	280.49	313.90	262.02
Lestrade's body	0.00	150.00	0.00

It can be seen from the data in Table 6.15 that, in the absence of subtitles, the target audience concentrated their full attention on the AOIs around Sherlock's and Lestrade's face, completely neglecting the AOIs around their bodies. The mean fixation durations in Table 6.16 largely confirm these fixation values, as on average some of the viewers looked at these AOIs the longest. This viewing behavior is linked to the camera distance, that is, the on-screen characters were filmed in medium close-up shots and thus only the upper part of their body was visible. Another reason why this AOI attracted so few fixations could be due to the fact that viewers tend to automatically reorient their gaze towards speaking faces, as it has also been observed to be the case in this study, especially when the voice modulation of the on-screen characters is of paramount importance to relay Sherlock's dissociative attitude. In addition, medium close-up shots enabled the participants to perceive a wide array of non-verbal resources pertaining to mise-en-scène, such as props (e.g., Sherlock's sunglasses), gaze direction (e.g., Sherlock looking at Lestrade from above) or facial expressions (e.g., Sherlock's smirking) that have been found to contribute to multimodal irony in this scene.

Contrary to my expectations, however, apart from the on-screen characters' faces, the participants did not explore other parts of the screen. This viewing behaviour appears to be part of a wider gaze pattern under non-subtitle conditions in which "viewers rarely explore the periphery of the image" and remain focused on the protagonists' faces (Smith, 2015). In the voiced-over clip, viewers also predominantly stick to the on-screen character's faces as in non-subtitle conditions. Here, these visual actions can be related to "a double soundtrack" that viewers have to deal with, that is, the original soundtrack and the voice-over narration, which in places may be confusing for audiences, especially when the acoustic modes (e.g., Sherlock's voice modulation) play a pivotal role in the construal of multimodal irony. For instance, the voice-over has a tendency to reduce viewers' abilities to recognise emotions and feelings through a limited access to intonation, speech rate or pitch because of the overlap between the narrator's voice and the character's voice. This assumption will be tested in the participants' responses to the irony comprehension questionnaire in the next chapter.

6.3.2 Analysis of clip: Sherlock and Lord Coward

SH1_3 is another case in which the acoustic modes play a pivotal role in the construal and comprehension of multimodal irony. There is also a whole range of non-verbal visual resources which highly contribute to the detection of Sherlock's mocking attitude. Counterintuitively, however, a number of viewers spent more time looking at the bottom than at the centre of the screen in the subtitled version. On average, as shown in Tables 6.17 and 6.18, the focus of attention switches from reading the subtitles towards processing the middle of the screen and it is immediately visible in all subgroups. The AOI around subtitles and the AOI around Sherlock's face competed for viewers' visual attention, while other AOIs remained largely ignored.

Table 6.17 Number of fixations and total number of fixations with the percent ratios in the AOIs in SH1_3 depending on the level of English proficiency (high, medium, low) in the SUB_VO group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Subtitles	39	36	42	39	61	56
Sherlock's body	16	15	13	12	10	9
Sherlock's face	40	37	41	38	25	23
Coward's body	10	9	10	9	10	9
Coward's face	2	2	1	1	2	2
Sum	107	100%	107	100%	108	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Subtitles	7249	27	8081	35	9923	43
Sherlock's body	5321	20	3047	13	1972	8
Sherlock's face	11971	45	9957	43	9171	39
Coward's body	1731	7	1605	7	1797	8
Coward's face	249	1	208	1	407	2
Sum	26521	100%	22898	100%	23270	100%

Table 6.18 Mean fixation duration in the AOIs in SH1_3 depending on the level of English proficiency (high, medium and low) in the SUB_VO group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Subtitles	185.87	187.93	165.38
Sherlock's face	317.65	263.42	246.84
Sherlock's body	332.56	234.38	197.20
Coward's face	124.50	208.00	203.50
Coward's body	173.10	160.50	179.70

From Table 6.17 we can see that HLPs, MLPs, LLPs had to rely on the subtitled text to some extent. Here, LLPs looked at the subtitles more frequently (61 fixations) and for longer periods (9923) ms respectively, than HLPs and MLPs, which suggests that LLPs had to activate more cognitive efforts to process the intended meaning in the subtitled text. High fixation values can be connected with their low level of English competence as they attempt to derive the implicit meaning from the verbal mode (the written text) which is the only one fully accessible to them. Surprisingly, HLPs and MLPs also had a high number of fixations on the subtitle area and spent about a quarter of their time reading the subtitles, as the total fixation durations indicate (Table 6.17), although the original dialogue was rendered with high-frequency words into one-line subtitles, and hence should be easier to process (e.g., Dwyer, 2015). The fact that HLPs and MLPs on average read the subtitles longer than LLPs is somehow surprising, as shown in Table 6.18. HLPs and MLPs' greater focus on reading

the subtitles than expected can be interpreted as an attempt to verify the source text against the target text translation, as a result of their high level of English proficiency. Thus, the missing question tag ("Is there?") in the subtitled text that has been identified as a key narrative cue to enhance the ironic effect in the Polish version might also draw their additional attention to the AOI around subtitles. On a different note, higher fixation values in the subtitled text can be also associated with their position of the subtitles in the part of the screen right under Sherlock's face as demonstrated in Figure 6.4.

Figure 6.4 Heat maps for HLPs, MLPs and LLPs in SH1_3 at 00:02:24

| HLPs | MLPs | LLPs |
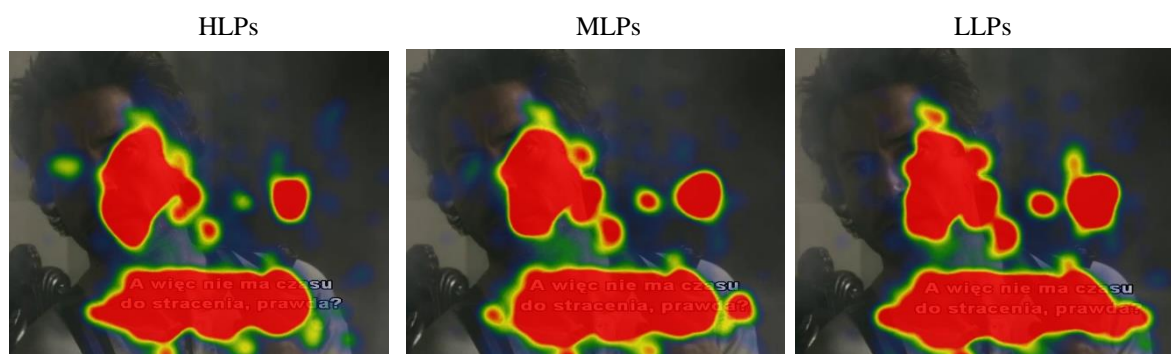
As presented in the visual representation in Figure 6.4, despite their reliance on the written captions, all participants managed to divide their attention between reading the subtitles and examining Sherlock's face, which again became the centre of attention for narrative and attentional cues. Considering the importance of non-verbal acoustic semiotic resources in this scene, Sherlock's voice quality may have attracted viewers' attention to his face, which would account for longer total fixation durations. The fact that Sherlock's face was displayed centrally on the screen in a medium close-up shot also gave many viewers the opportunity to grasp more detailed traits and features of Sherlock's facial features that were identified as key non-verbal resources construing multimodal irony in this clip (see Section 5.2), and thus contribute to their emotion recognition process. This expectation, however, will be verified in the analysis of the questionnaire responses in the next chapter. As noted by Sita et al. (2015) and shown in Figure 6.5 above these back and forth fixations between the eyes, mouth and face suggest that "viewers are engaging with the scene as they would do in a normal face-to-face encounter, using eye movements to verify who people are and what they are feeling", despite the presence of subtitles at the bottom of the screen.

Let us now consider the analysis of eye movements in the voiced-over version. As shown in Table 6.19, Sherlock's face again remained a great gaze attractor, while other AOIs were overlooked by HLPs, MLPs and LLPs.

Table 6.19 Number of fixations and total fixation duration with the percent ratios in the AOIs in SH1_3 depending the level of English proficiency (high, medium, low) in the VO_SUB group

| Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	2	1	9	10	8	6
Sherlock's face	139	95	77	85	114	92
Coward's body	4	3	2	2	0	0
Coward's face	2	1	3	3	2	2
Sum	147	100%	91	100%	124	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Sherlock's body	917	2	2080	5	2841	6
Sherlock's face	36541	91	38079	93	41202	92
Coward's body	1783	4	433	1	0	0
Coward's face	750	2	541	1	733	2
Sum	39991	100%	41133	100%	44776	100%

Table 6.20 Mean fixation duration in the AOIs in SH1_3 depending on the level of English proficiency (high, medium and low) in the VO_SUB group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Sherlock's face	263.51	431.20	361.42
Sherlock's body	458.50	208.00	355.13
Coward's face	375.00	180.33	366.50
Coward's body	445.75	216.50	0.00

The results obtained from the analysis of eye-tracking data in Table 6.19 show a similar viewing pattern for all participants who found the AOI around Sherlock's face definitely the most significant out of all AOIs. Since the AOI around Sherlock's face was centrally located in a close-up shot, all viewers could easily explore specific features of his facial expressions. Their eye movements moving back and forth between the eyes and the mouth additionally indicate that viewers were able to conduct a detailed emotional and facial examination to support their retrieval of the intended meaning. The fixation values on Sherlock's face can be also linked to the voice-over narration. Since the film score and Sherlock's voice modulation have been identified as two of the main non-verbal resources construing multimodal irony in SH1_3, the voice-over narrator hiding a great majority of the original soundtrack could make it more effortful for some participants to process the target and source dialogue simultaneously. As a result, the exposure to the voice-over narration could also account for longer total fixation duration and greater number of fixations to allow viewers to recognise Sherlock's emotions, and thus interpret his attitude of dissociation. It is also

163

interesting to note that the viewers did not take the opportunity to explore other AOIs despite the absence of subtitles. This may indicate that none of the remaining AOIs in the middle of the screen were considered particularly relevant to the narrative of the scene. Nevertheless, HLPs, MLPs and LLPs shifted their attention to other elements of mise-en-scène outside of the delineated AOIs, as visible in the heat maps in Figure 6.5.

Figure 6.5 Heat maps for HLPs, MLPs and LLPs in SH1_3 at 00:02:23

HLPs	MLPs	LLPs
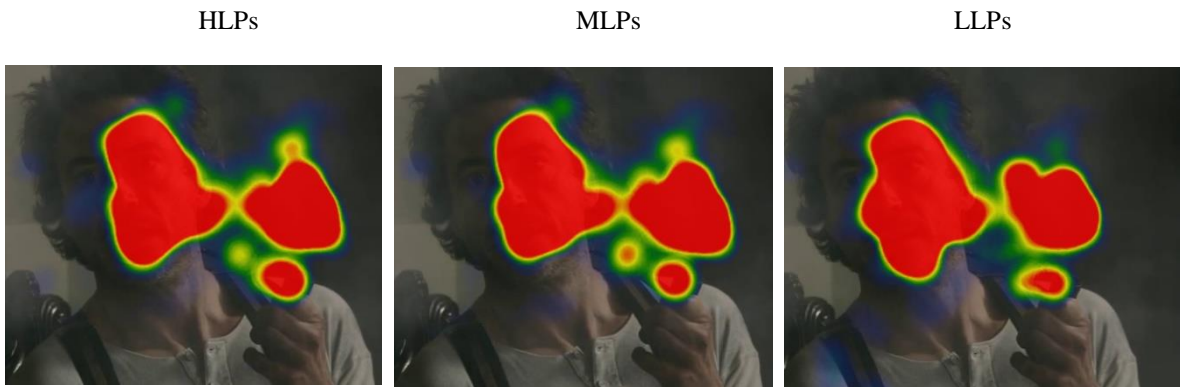
As can be seen in Figure 6.6, for instance, that several subjects noticed a pipe, which was also identified as an object of narrative importance in the construal of multimodal irony (see Section 5.2), although it was not a part of any AOIs. Surprisingly, however, as shown in Table 6.20, the mean fixation durations illustrate that on average the participants dwelled on Sherlock's face relatively long but also these few viewers who spotted the AOIs around Coward's face and body stayed focused on these areas of interest for a significant amount of time.

6.3.3 Analysis of clip: Watson's stag party

In the subtitled version of **SH2_2,** the AOIs around faces of the on-screen characters were again the most visually attractive areas of the screen, as expected. These AOIs drew and captivated viewers' attention for much longer then the AOI around subtitles, as indicated in Table 6.21. Surprisingly, even LLPs spent more time on the AOIs in the middle of the screen than on the subtitle area, which potentially increased their chances to identify several non-verbal visual resources in addition to the information conveyed acoustically.

Table 6.21 Number of fixations and total fixation duration with the percent ratios in the AOIs in SH2_2 depending on the levels of English proficiency (high, medium, low) in the VO_SUB group

| Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	1	1	0	0	0	0
Watson's face	74	47	66	40	53	35
Sherlock's body	0	0	1	0	0	0
Sherlock's face	36	23	42	25	39	26
Mycroft's body	0	0	1	1	3	2
Mycroft's face	28	18	30	18	34	22
Subtitles	19	12	25	15	23	15
Sum	158	100%	165	100%	152	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	158	0	0	0	0	0
Watson's face	19651	49	19643	47	18688	45
Sherlock's body	0	0	0	0	0	0
Sherlock's face	7760	19	8748	21	7974	19
Mycroft's body	0	0	166	0	549	1
Mycroft's face	9648	24	8487	20	9979	24
Subtitles	3259	8	4451	11	4185	10
Sum	40476	100%	41495	100%	41375	100%

Table 6.22 Mean fixation duration in the AOIs in SH2_2 depending on the level of English proficiency (high, medium and low) in the VO_SUB group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Subtitles	171.53	178.04	180.78
Watson's face	278.34	297.62	352.60
Watson's body	158.00	0.00	0.00
Sherlock's face	215.56	208.29	204.46
Sherlock's body	0.00	166.00	0.00
Mycroft's face	345.86	282.90	293.50
Mycroft's body	0.00	166.00	183.00

As can be seen from Table 6.21, HLPs, MLPs and LLPs clearly avoided gaying at the captions to a great extent mirroring the irregular way of reading these one-line subtitles as found in the literature (see d'Ydewalle and de Bruycker, 2007). Since all viewers were reading the subtitles for much shorter than exploring other AOIs in the middle of the screen, i.e., 3259 ms (HLPs), 4451 ms (MLPs) and 4185 ms (LLPs), they probably did not need to activate their mental processes to a great extent to comprehend the intended meaning. This viewing behaviour is also relevant for LLPs who were expected to spend more time reading the captions, as a result of their low level of English proficiency. It is likely that these gaze

patterns were therefore determined by the subtitle translation strategy used as well as the quality of words selected to render the original dialogue in Polish. That is to say, a verbatim translation as well as the use of high-frequency words made the subtitled text easier to comprehend, despite the linguistic complexity of the intended meaning. This assumption will be verified in the following chapter. However, subtitles on their own do not convey enough information to realise that Mycroft's dialogue was not intended by filmmakers to be understood literally. Although viewers usually reorient their gaze towards speaking faces on the screen (e.g., Kruger, Szarkowska and Krejtz 2015). It can be also confirmed by the mean fixation durations in Table 6.21 illustrating that on average all participants were predominantly fixating on the AOIs around the on-screen character's faces. In this scene they explored the AOI around Watson's face the longest time. This viewing movements can be informed by Watson's "antiphonal laughter" (Bryant, 2011) that determined the construal of ironic meaning. While laugher is considered the most powerful marker of ironic content (Jefferson, 1984), it also appears to draw visual attention in this clip from the participants considerably. Here the visual and kinesic modes also come into play. The importance of the AOIs in the middle of the screen is confirmed by the number and total duration of fixations on the on-screen character's faces, as shown in Table 6.21, which convey visually salient pieces of information in the construal of multimodal irony. As seen from the analysis of data in Table 6.21, despite the presence of subtitles in the bottom of the screen, many viewers could split their attention between the subtitles and the visual action in the middle of the screen rather effortlessly irrespective of their level of English language proficiency.

With regard to the voiced-over version, several participants looked at the AOIs in a very similar manner. That is, their eyes predominantly remained still in the centre of the screen examining the on-screen characters' faces the most frequently and the longest, as illustrated in Table 6.23.

Table 6.23 Number of fixations and total fixation duration with the percent ratios in the the AOIs in SH2_2 depending on the level of English proficiency (high, medium, low) in the SUB_VO group

Measurement	Fixation count	Percent ratio	Fixation count	Percent ratio	Fixation count	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	0	0	0	0	2	2
Watson's face	61	39	55	41	35	35
Sherlock's body	0	0	0	0	0	0
Sherlock's face	50	32	44	33	30	30
Mycroft's body	0	0	0	0	3	3
Mycroft's face	44	28	34	26	31	31
Sum	155	100%	133	100%	101	100%
Measurement	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio	Total fixation duration (ms)	Percent ratio
AOIs	HLPs		MLPs		LLPs	
Watson's body	0	0	0	0	225	1
Watson's face	16787	44	15599	46	12446	44
Sherlock's body	0	0	0	0	0	0
Sherlock' face	9397	25	8271	25	6123	22
Mycroft's body	0	0	0	0	299	1
Mycroft's face	12093	32	9887	29	9148	32
Sum	38277	100%	33757	100%	28241	100%

Table 6.24 Mean fixation duration in the AOIs in SH2_2 depending on the level of English proficiency (high, medium and low) in the SUB_VO group

Measurement	Mean fixation duration		
AOIs	HLPs	MLPs	LLPs
Watson's face	275.20	283.62	355.60
Watson's body	0.00	0.00	112.50
Sherlock's face	187.94	187.98	204.10
Sherlock's body	0.00	0.00	0.00
Mycroft's face	274.84	290.79	295.10
Mycroft's body	0.00	0.00	99.67

It is apparent from Table 6.23 that there are a number of similarities in the percent ratios for the number and total fixation durations among HLPs, MLPs and LLPs who looked at the on-screen characters' faces in a more consistent way. This viewing behaviour suggests that participants possibly wanted to identify the role that each on-screen character plays in the construal of intended meaning in this scene. Although many participants tended to spend more time looking at Watson's face, they also explored Sherlock's and Mycroft's faces to a greater extent than the viewers watching the scene with subtitles (Table 6.22), as confirmed by the mean fixation durations in Table 6.24. Hence, Watson's "antiphonal laughter" (Bryant, 2011) once more turned out to be a strong attractor of visual attention which should be clearly audible because of the absence of voice-over narration in this particular fragment

of the scene. Voice-over, however, reduced the audibility the majority of Mycroft's dialogue, drowning out his characteristic voice qualities. Instead, participants explored Mycroft's facial expressions, body and head position or camera angle that were also expected to support their retrieval of Mycroft's dissociative attitude towards Watson.

6.4 Tests for Statistical Significance

The analyses of the eye-tracking data in the clips above confirmed the predictions that I had made in Section 6.1. That is, the number of fixations, the total fixation durations as well as the mean fixation durations showed that HLPs and MLPs looked generally less at the subtitles and more at the AOIs around the on-screen characters' faces, as anticipated, although the proportion of the time spent on these AOIs varies from clip 1 to 6. LLPs, on the other hand, read the subtitles for longer than HLPs and MLPs, as expected. Nevertheless, they also managed to dwell on the AOIs around the faces of the on-screen characters.

Now, I attempted to see whether these fixation values are statistically significant for these language subgroups. Because of the small number of participants in each group (SUB_VO and VO_SUB), and hence relatively few eye-movement data, it was not possible to test the difference in each clip separately. As a result, statistical tests were carried out in R Studio on the basis of the mean fixation durations calculated across all clips together in SUB_VO and VO_SUB groups for HLPs, MLPs and LLPs. Prior to the selection of a statistical test, a normality test was conducted to verify whether the data is normally distributed or not. In Figures 6.6 and 6.7 below the histogram and Q-Q plot illustrate the data distribution in the current eye-tracking experiment.

Figure 6.6 Histogram presenting the data distribution
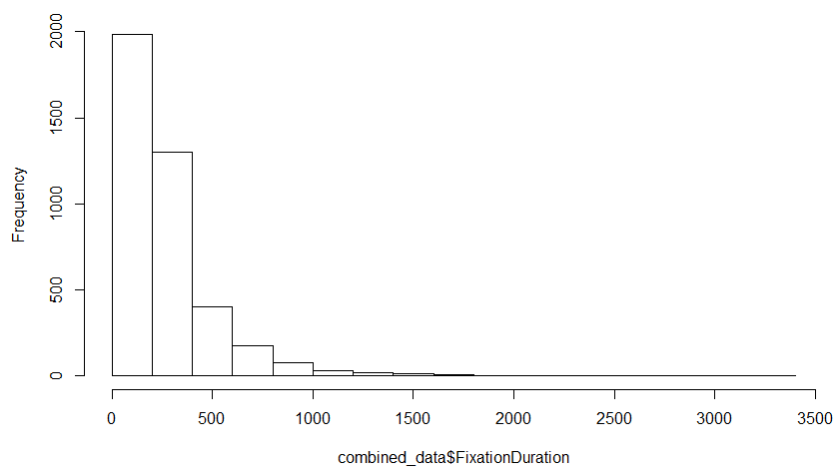
Figure 6.7 Q-Q plot presenting the data distribution
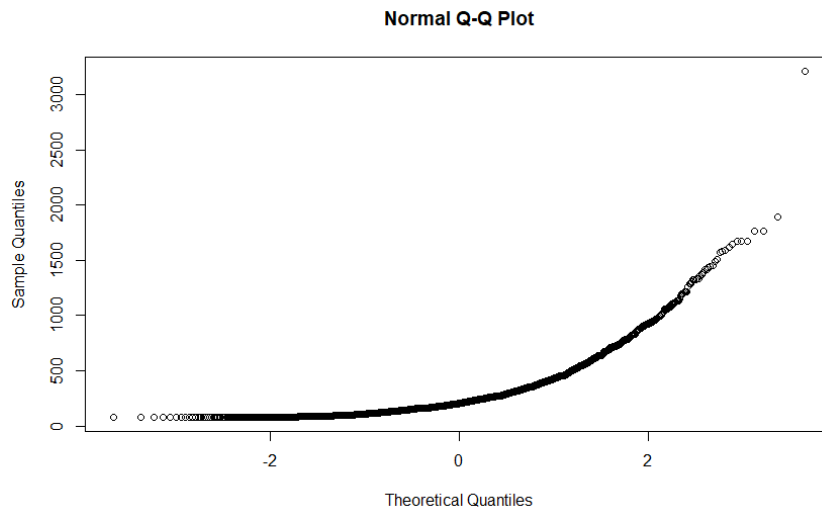
As can be seen in Figures 6.6 and 6.7, the histogram and QQ plot show us visually that the data is not normally distributed. This is also confirmed by a highly significant Shapiro-Wilk normality test ($W = 0.733$, $p < .001$), as demonstrated in Figure 6.8 below.

Figure 6.8 Results retrieved from the Shapiro-Wilk normality test

The distribution of the fixation duration data can also be visualised for each of the language subgroups for the main groups (SUB_VO and VO_SUB) showing the spread of data for the subgroups in Figure 6.9 below.

Figure 6.9 Visualisation of the spread of data for HLPs, MLPs and LLPs subgroups in SUB_VO and VO_SUB groups

169
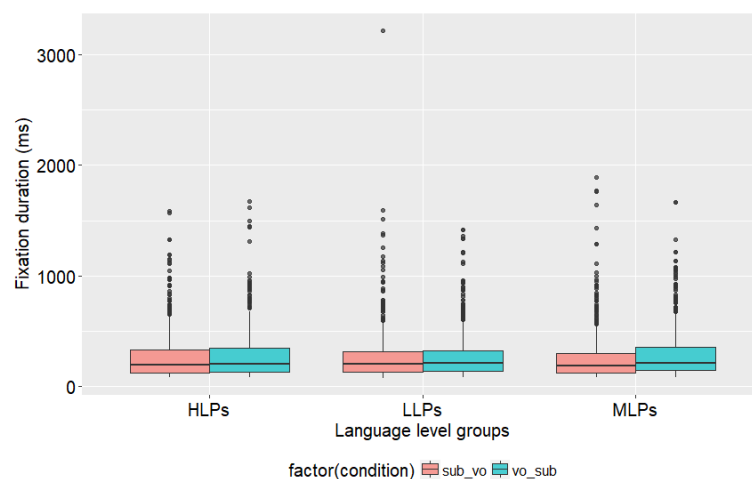
As the data is non-normally distributed I had to use non-parametric tests to analyse the data. To do this I used Mann-Whitney U tests to look for differences between each of the sub groups due to the un-equal group sizes. To account for the increase in Type I error (i.e., the more tests we run the more likely we are to detect a significant result by chance) I applied a continuity correction. The results retrieved from Mann-Whitney U test are reported in Figure 6.10 below.

Figure 6.10 Mann-Whitney U test correlation tests for HLPs, MLPs and LLPs in SUB_VO and VO_SUB

```
        Wilcoxon rank sum test with continuity correction

data:  MLPs_sub_vo$FixationDuration and MLPs_vo_sub$FixationDuration
W = 246700, p-value = 0.0005091
alternative hypothesis: true location shift is not equal to 0
```

```
        Wilcoxon rank sum test with continuity correction

data:  HLPs_sub_vo$FixationDuration and HLPs_vo_sub$FixationDuration
W = 199470, p-value = 0.2074
alternative hypothesis: true location shift is not equal to 0
```

```
        Wilcoxon rank sum test with continuity correction

data:  LLPs_sub_vo$FixationDuration and LLPs_vo_sub$FixationDuration
W = 177220, p-value = 0.1269
alternative hypothesis: true location shift is not equal to 0
```

	sub_vo	vo_sub		
Language level	M (SD)	M (SD)	W	P
LLPs	273 (245)	281 (213)	177220	0.126
MLPs	260 (223)	286 (222)	246700	0.0005*
HLPs	271 (217)	285 (226)	1994470	0.2074
Note: * = p < 0.016				

As revealed in Figure 6.10, a set of Mann-Whitney U tests were carried out on the fixation durations for the language subgroups using Bonferroni correction ($\alpha = 0.016$) to see if there is any significant difference. There were no significant differences for the language subgroups, with the exception of MLP, which highlighted 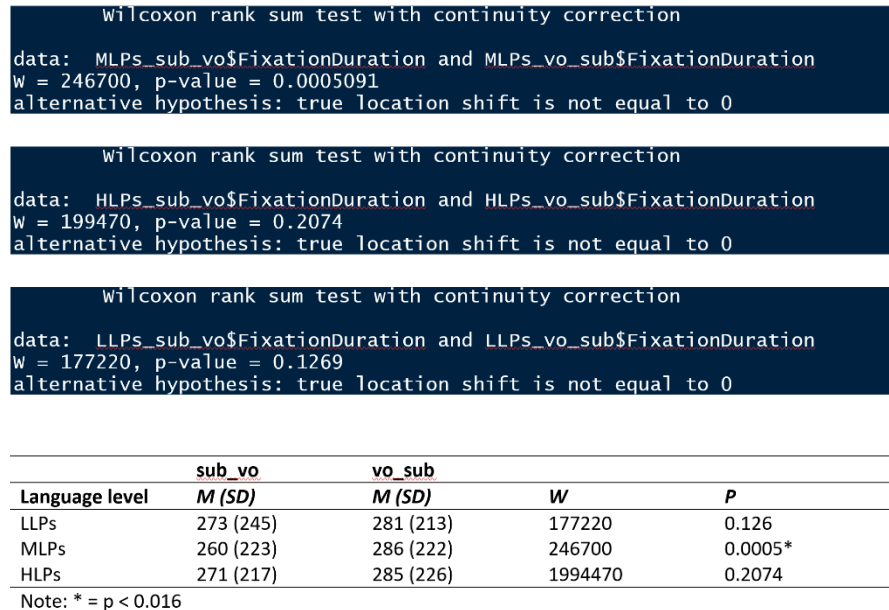a significant difference between the 2 groups (SUB_VO and VO_SUB) for fixation duration ($W = 246700$, $p < .016$). In addition, the Pearson's correlation test was conducted to see whether there is any statistical relationship between fixation duration and different language level, as illustrated in Figure 6.11.

Figure 6.11 Pearson's correlation test for HLPs, MLPs and LLPs in SUB_VO and VO_SUB groups

```
        Pearson's product-moment correlation

data:  combined_data$FixationDuration and combined_data$group
t = 0.055313, df = 4002, p-value = 0.9559
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.03010243  0.03184945
sample estimates:
        cor
0.0008743487
```
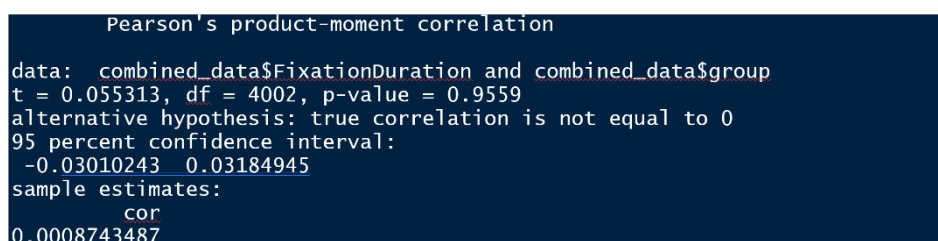
As presented in Figure 6.11 there is no overall correlation between fixation duration and language level ($t(4002) = 0.055$, $p = .955$, $r = 0.00087$). Although, the statistical tests did not reveal statistically significant data, which can also be influenced by a relatively low number of participants in each subgroup, the presented analyses of the eye-tracking data constitute a major step towards our understanding of viewing behaviour of the viewers at different levels of proficiency in English and the way how they engage with subtitled and voiced-over clips featuring the presence of multimodal irony.

6.5 Conclusions

The main goal of this chapter was to gain better understanding of how HLPs, MLPs and LLPs distributed their visual attention on the screen across AOIs to interpret the attitude of dissociation, and thus the intended ironic meaning behind it, in the subtitled and voiced-over clips of *Sherlock Holmes* (2009) and *Sherlock Holmes: A Game of Shadows* (2011).

Based on the analysis of the eye-tracking data in the subtitled clips, it is evident that all viewers exhibited notable differences in viewing behaviour in respect to the level of their abilities to divide attention among different AOIs on the screen. Their different levels of English proficiency have been shown to be a principal factor determining the way in which attention was split between reading the subtitles and processing the visual action in the centre similarly to the visual patterns found in other studies (e.g., Orrego-Carmona, 2015; Szarkowska et al., 2016). For instance, the majority of HLPs and MLPs often skipped the subtitles – they tended to scan them, rather than fully read them – allowing themselves a greater proportion of time to extract information from other AOIs in the middle of the screen. These gaze patterns could be linked to their high level of familiarity with the foreign language and culture, I argue, as it appears that the processing of the emotion words, idiomatic expressions or semantically significant phrases to the narrative was less challenging for them. This is confirmed by shorter total fixation duration in the subtitle area among these groups reflecting the relatively small amount of metal effort required to process the instances of multimodal irony under analysis. LLPs, on the other hand, tended to look at the AOIs around subtitles more frequently and for longer periods than HLPs and MLPs. These eye movements suggest that LLPs needed more time to process either phrases of importance to the narrative or emotion words in order to retrieve the intended meaning. These gaze patterns indicate that these participants experienced some difficulties switching attention among multiple AOIs, behavior, which is associated with a high cognitive load. On some occasions, however, LLPs spent a comparable amount of time reading the subtitles as

other viewers did. Despite HLPs, MLPs and LLPs' reliance on the subtitles, to a various extent, and the additional cognitive load imposed on their visual perception system, they managed to explore other AOIs in the middle of the screen predominantly focusing on the on-screen characters' faces. Thus, the display of subtitles in the bottom of the screen did not constrain their access to the non-verbal visual resources. LLPs also searched for narrative cues in a more explorative way looking at each AOI in similar proportions and dwelling on the visually peripheral resources outside of the AOIs, too.

In the voiced-over clips, HLPs, MLPs and LLPs exhibited very similar gaze patterns when looking at the AOIs located centrally on the screen, irrespective of the level of proficiency in English, as anticipated. They predominantly looked at the middle of the screen prioritising the AOIs around the on-screen characters' faces. Surprisingly, some subjects hardly looked at other AOIs, largely ignoring the AOIs around their bodies, for instance, although they had more chances to share their attention across other parts of the screen in the absence of subtitles. Nevertheless, the fixation values on the AOIs around faces are only slightly higher in the voiced-over than in the subtitled version, which suggests that many viewers could identify a great deal of the non-verbal visual semiotic resources irrespective of the modality of translation used to relay multimodal irony.

As it has also been the case in the previous studies (e.g., Treuting, 2006, Sita et al., 2015), the AOIs around faces of the on-screen characters have also been considered to be the strongest attractors of visual attention in the context of this thesis on multimodal irony. These results underline the semiotic significance of facial expressions (e.g., smirks, wide open eyes, raised eyebrows) as a source of the most meaningful non-verbal information in the detection of the attitude of dissociation. Thus, the AOIs around bodies of the on-screen characters were not of particular interest to Polish viewers, unless these relayed salient visual cues relevant to the narrative, as it has been found in SH1_2. As a result, in the multimodal texts like films, verbal and non-verbal resources have to compete for viewers' attention and this competition is often based on "the prioritisation of information" principle (Kruger, 2012: 83) to distinguish elements with high visual saliency that also have high narrative saliency (see Kruger, 2012).

Last but not least, the findings are particularly important in the light of the audiovisual scene in Poland where voice-over is still a dominant audiovisual translation modality on TV. Although the perception of non-verbal narrative cues may be easier in the voiced-over version of the clips under analysis due to the absence of subtitles, it does not mean that all viewers will grasp the intended meaning more successfully. In the subtitled

version, the results showed that participants with high and medium level of English switched their attention across several AOIs rather effortlessly. Even LLPs managed to explore the AOIs in the middle of the screen to a certain extent, despite the display of subtitles at the bottom. It thus appears that subtitles could be included as an optional type of translation on TV in Poland. Before any further analysis and recommendation, though, it is vital to examine whether the target audience actually retrieved the intended ironic meaning in the subtitled and voiced-over excerpts under scrutiny.

The chapter that follows presents the results obtained from the questionnaire in the interactionist component of this study in order to supplement the eye-tracking data. The analysis of questionnaire data aims to verify whether and/or to what extent the participants in this experiment perceived and understood the attitude of dissociation, and thus the ironic meaning behind it. It also aims to identify some characteristics of eye movements in the certain AOIs that are related to reading comprehension skills in the multimodal texts, as previously highlighted in other studies (Schmidt-Weigand et al., 2010).

7 INTERACTIONIST COMPONENT: QUESTIONNAIRE

7.1 Introduction

In the previous chapter, the eye-tracking data collected from HLPs, MLPs and LLPs watching the subtitled and voiced-over clips of SH1 and SH2 have been analysed. These results revealed that their viewing behaviour is to some extent determined by their level of English proficiency and the type of translation that multimodal irony is relayed with in the TL. While the participants who viewed the voiced-over version remained focused mainly on the on-screen character's face, subjects viewing its subtitled counterpart divided their visual attention between the bottom and the middle of the screen, to various extents, to read the subtitles and follow the action in the clips, respectively. On several occasions, it has been found that HLPs, MLPs and in particular LLPs read the subtitled text slower. Since my analysis of the eye-tracking data is informed by the premise that "the longer the fixation time [...], the greater the cognitive effort" required to process the fragment of the clip in question (Saldanha and O'Brien, 2013: 144), the gaze patterns of HLPs, MLPs and LLPs have been accounted for in terms of cognitive effort that is required to understand a more complex concept like irony.

Chapter 7 reports on the last strand of my analysis. The interactionist component has been used as a tool to measure whether and/or to what extent HLPs, MLPs and LLPs retrieved meaning and, more importantly, ironic meaning in the subtitled and voiced-over excerpts under analysis. In this chapter, the eye-tracking data have thus been complemented with the analysis of the participants' responses to a questionnaire probing their irony comprehension. This questionnaire-based strand of my analysis seeks to verify whether longer fixations on subtitles, and thus a slower reading of them, hinder the perception of visual information and impede comprehension of the intended ironic meaning; such verification involves participants with different degrees of proficiency in the English language and familiarity with the British culture. It has also been examined whether longer fixation durations are also linked to other variables like the layout of subtitles in specific parts of the scene. Analogously, it is intended to unveil whether longer fixations on the on-screen character's faces allowed the participants to process more non-verbal visual semiotic resources, and thus enhance the retrieval of the intended meaning, or whether some viewers took longer to explore the AOIs around faces to grasp acoustic information that is particularly salient in the original clip narrative, but more difficult to access through the

voice-over narration. Additionally, in this chapter, it has been examined whether the participants grasped the non-verbal semiotic resources that have been considered vital in the construal of multimodal irony in the descriptive component, and to what extent these non-verbal cues are recognised across cultures.

The comprehension of the instances of multimodal irony that were composed primarily through visual and kinesic modes is examined in Section 7.2, and through acoustic modes in Section 7.3 with the attempt to reveal the extent to which the verbal and non-verbal modes contributed to the retrieval of the intended meaning. In Section 7.4, it is discussed whether certain non-verbal resources construing the dissociative attitude on screen in the original text are recognised as culturally specific, and thus whether they are more challenging to interpret for the target audience. This chapter ends with Section 7.5, in which the findings obtained from the analysis of the questionnaire data are summarised. For each response, each viewer was assigned individual scores (see Appendix 8 for the participants' responses).

7.2 Reception of Visually Construed Multimodal Irony

Section 7.2 examines the responses given by the viewers in the questionnaire about their comprehension of the subtitled and voiced-over clips of SH1 and SH2. In these excerpts, the visual and kinetic modes have been recognised as the main resources used in the construal of Sherlock's or Watson's dissociative attitude (i.e., SH1_2; SH2_1; SH2_3), with the acoustic mode identified as supportive. In the following pages, it is discussed how respondents with different levels of English knowledge, i.e., high, medium and low level, grasped the attitude of disapproval.

7.2.1 Reception of clip: Watson and Sherlock

In **SH1_2**, as anticipated, the vast majority of the participants in SUB_VO and VO_SUB groups correctly understood on the whole Watson's attitude of dissociation and his ironic intentions towards Sherlock. Table 7.1 displays the mean comprehension scores (see Subsection 4.6.3) that HLPs, MLPs and LLPs earned for their responses to the open-ended questions (Q1 and Q2). On average, HLPs and MLPs retrieved multimodal irony successfully when watching the excerpts with subtitles and voice-over. Counterintuitively, LLPs received much fewer comprehension points in the subtitled clip and scored only marginally higher in the voiced-over clip.

Table 7.1 Mean comprehension scores of HLPs, MLPs and LLPs in the subtitled (SUB_VO) and voiced-over (VO_SUB) versions (Appendix 8)

Mean	HLPs	MLPs	LLPs
Subtitles SUB_VO	5.4	7.2	3.5
Voice-over VO_SUB	5.6	4.8	3.8

As illustrated in Table 7.1, HLPs grasped the ironic meaning both in the subtitled (mean=5.4) and in the voiced-over (mean=5.6) version rather easily, as expected. In response to Q1, a great number of HLPs explained precisely why Watson is speaking to Sherlock the way he does (e.g., P01A, P07A, P19A), describing his attitude of dissociation correctly as mocking or sarcastic (e.g., P08A, P019A, P11B) as well as "żartobliwy, ironiczny" (BT: "jocular, ironic") (e.g., P01A) in Q2. Their mean comprehension scores also suggest that neither the subtitles nor the voice-over hampered their retrieval of the intended meaning in this scene. Similarly, MLPs also interpreted the non-literal meaning of Watson's statement rather effortlessly collecting even more points in irony comprehension in the subtitled version (mean=7.2) than in its voiced-over (mean=4.8) counterpart, as Table 7.1 demonstrates. As shown by the eye-tracking data in Subsection 6.2.1, HLPs and MLPs largely ignored looking at the subtitles since they did not have to rely on them to retrieve the intended meaning most likely as a result of their high proficiency in English. Instead, HLPs and MLPs listened to the original dialogue directly and focused their attention fully on the centre of the screen. Since HLPs and MLPs predominantly dwelled on the AOIs around Watson's face and Sherlock's body in SUB_VO and VO_SUB groups, they were able to spot a great number of non-verbal visual cues construing multimodal irony on screen. A solid evidence base is provided in their responses to Q3. For instance, virtually all HLPs and MLPs made interesting and detailed observations pointing out non-verbal semiotic resources pertaining to the visual and kinesic modes of film language. To name just a few, several subjects singled out Sherlock's terrible appearance, garment, dark room's decorations, lighting, objects as well as Watson's gestures, head movements, nonchalant and straight posture and mimics and, in particular, his smirking both in the subtitled and voiced-over excerpts (e.g., P07A; P08A; P16A; P18A). In addition to the original soundtrack, HLPs and MLPs predominantly relied on the information conveyed visually to successfully capture the emotional bond between Watson and Sherlock. They also depended on the visual cues to identify the incongruity between Watson's ironic riposte and Sherlock's looks, although they struggled

to grasp Watson's intonation in the voiced-over clip due to the voice-over narration, as indicated in Q4. As a result, Poles with high or medium level of proficiency in English are more likely to recognise some culture-specific emotions, as confirmed in previous studies claiming that "L2 proficiency […] had a strong effect on the perception of emotions in the L2" (Ożańska-Ponikwia, 2013: 14).

On the other hand, several participants with low level of English overall misinterpreted the filmmakers' intentions, which caused a serious loss in irony comprehension in Clip SH1_2. That is to say, in contrast to HLPs and MLPs, LLPs earned fewer points both in the subtitled clip (mean=3.5) and its voiced-over (mean=3.8) counterpart. This lack of comprehension is demonstrated in the incorrect responses that some of the viewers (e.g., 13A) gave to Q1. For instance, the participant 13A said that "Sherlock wygląda niekorzystnie. Przyjaciel zwraca mu uwagę, że wygląda niewłaściwie" (BT: "Sherlock looks unfavourably. A friend remarked that his look is inappropriate"). This misinterpretation is further verified in the responses to Q2, in which some of those surveyed described Watson's attitude towards Sherlock as "po przyjacielsku, z troską" (BT: "friendly, concerned") (P13A) or "dość ciepły" (BT: "warm/hearty") (P21A), which is the opposite to the intentions that the filmmakers wanted to communicate with Watsons' statement. These inadequate descriptions, for instance in the subtitled clip, are not caused by the participants' limited access to the non-verbal visual semiotic resources in the middle of the screen due to the display of subtitles. As shown in Subsection 6.2.1, LLPs spent about 20% of their time looking at the subtitles area; during the rest of the time, they mainly fixated on the AOIs around Watson's face and Sherlock's body. In fact, as indicated in responses to Q3, LLPs spotted some of the non-verbal resources pertaining to mise-en-scène. For example, 5 out of 8 LLPs in the SUB_VO group mentioned the dark scenery, Sherlock's looks or Watson's facial expressions in the subtitled clip (e.g., P22B; P12A; P03A). In the voiced-over clip, against my expectations, only 4 out of 6 LLPs in VO_SUB group identified some pieces of information relayed visually, although they had full access to the centre of the screen in the absence of subtitles (e.g., P13B; P21B). Nevertheless, LLPs in both groups provided more general descriptions of visual cues and failed to grasp more detailed information (e.g., Watson's smirking) that could enhance their comprehension of Watson's ironic comment and the emotions linked to it. This is also supported by 2 LLPs in the SUB_VO group, who admitted in Q4 that subtitles dragged their eyes away from the middle of the screen, and thereby they could not focus on details (e.g., P03A, P30A). In the VO_SUB group, 5 LLPs, on the other hand, acknowledged that they struggled to grasp Watson's tone of voice or

intonation and, by extension, his emotions due to the narrator's voice interfering with the original dialogue. Apart from the obstacles posed by the subtitles and voice-over translation, LLPs' abilities to comprehend complex concepts like irony and emotions associated with it can be constrained by their limited understanding of the foreign language and culture. According to Ożańska-Ponikwia (2013: 8) apart from "possible linguistics misunderstandings, there is a cultural and emotional mismatch which could be overcome by contact with L2 culture" that LLPs are largely missing.

7.2.2 Reception of clip: Sherlock, Watson and a scarf

SH2_1 is the instance of multimodal irony that nearly all participants understood accurately in the selected excerpt with subtitles and voice-over, contrary to my expectations regarding LLPs in the SUB_VO group who also accessed the intended meaning in the subtitled clip without considerable problems. On the whole, the responses given by HLPs, MLPs and LLPs to Q1 and Q2 in this scene indicate that nearly all respondents derived the intended meaning relayed both with subtitles and voice-over almost equally, as Table 7.2 illustrates.

Table 7.2 Mean comprehension scores of HLPs, MLPs and LLPs in the subtitled (VO_SUB) and voiced-over (SUB_VO) versions (Appendix 8)

Mean	HLPs	MLPs	LLPs
Subtitles SUB_VO	5.4	7.2	3.5
Voice-over VO_SUB	5.6	4.8	3.8

It is apparent from Table 7.2 that the vast majority of the subjects recognised Sherlock's dissociative attitude as construed by the filmmakers, and Sherlock's ironic riposte as relayed with subtitles and voice-over by the audiovisual translators. High comprehension scores earned by HLPs, MLPs and LLPs are motivated by their accurate responses to Q1 and Q2, where the respondents precisely described Sherlock's attitude to the proposition echoed to mock Watson and his future wife. In their responses, many viewers also clearly articulated the reason why Sherlock mocks Watson the way he does (e.g., P01B; P10B; P10A; P06A; P04A, P20A). For example, P10A responded that Sherlock behaves towards Watson "ironicznie i lekceważąco" (BT: "ironically and disrespectfully"), while P06A responded "lekceważąco, z typową dla siebie wyższością" (BT: "disrespectfully with his typical sense of superiority").

Since HLPs read the subtitles for only 6946 ms, as Subsection 6.2.2 reveals, I assumed that they would retrieve the intended ironic meaning with little effort, which has been confirmed by their mean comprehension scores in Table 7.2. HLPs also supported their understanding of Sherlock's disapproving attitude with a combination of the non-verbal semiotic resources that they fixated on in the AOIs in the middle of the screen (see Subsection 6.2.2), as verified in their responses to Q3. HLPs described Sherlock's ironic behaviour in significant detail. For instance, they pointed to Sherlock's eyebrow movement, posture, serious facial expression, eye contact as well as distance between the two on-screen characters, among others, in the subtitled version (e.g., P19B, P06B); they also referred to Sherlock's voice qualities in its voiced-over counterpart (e.g., P07A, P10A, P19A). P19B's response can be used by the way of illustration: "Gesty, ruch reką, postawa jaka przybrał (jedna noga zgięta, nieco teatralnie), ruch brwi i na pewno poważna mina przy wypowiadniu tych słów oraz sceneria w której to Watson nie pasował to ogólnego obrazu" (BT: "Gestures, hand's movements, his posture (one leg bended, a bit theatrical), eyebrows movement, a serious face while speaking, settings to which Watson did not fit at all"); P10A's response is also useful in this regard: "mimika twarzy, możliwość usłyszenia przed lektorem tonu głosu bohatera, jego postawa" (BT: "mimics, the protagonists' tone of voice heard before the narrator, his posture").

Contrary to my expectations, MLPs and LLPs also achieved high comprehension scores in the subtitled and voiced-over excerpts. Although, MLPs and LLPs read the two-line subtitles much more slowly than HLPs as the eye-tracking data indicates, they still understood the phenomenon under study accurately. As a result, in this instance of multimodal irony, which was modified and paraphrased in the subtitles with high-frequency and colloquial words, longer fixations are not associated with greater cognitive load. Instead, they would appear to be motivated by the length of subtitles, which did not hamper the participants' abilities to retrieve the ironic meaning. Interestingly enough, despite the time MLPs and LLPs fixated on the subtitles, they still managed to spot several visual cues that are important to the narrative, as indicated in their responses to Q3. They distinguished Sherlock's mimics, body movements or gestures, among others (e.g., P19B; P16B; P01B; P17B; P22A, P09A), although neither MLPs nor LLPs provided the same level of detail in their observations of the non-verbal semiotic resources as HLPs did.

As far as the supportive role of the acoustic modes is concerned, 9 out of 21 participants watching the subtitled clip noted the contribution of Sherlock's voice modulation, including tone, intonation and volume, as well as the music score rather easily

(e.g., P06B; P20B; P18B; P21B; P13B). Similarly, 8 out of 23 viewers watching the voiced-over clip spotted some features of Sherlock's voice quality as well as the film score despite the narrator's voice hiding most of the original soundtrack.

7.2.3 Reception of clip: The ambush on a train

Consider now the excerpt **SH2_3**, in which a large number of HLPs and MLPs provided a substantial evidence base to explain how they had been able to identify Sherlock's mocking attitude towards Watson, both in the subtitled and voiced-over version, as shown in Table 7.3. Nevertheless, as anticipated, Sherlock's attitude posed an interpretative challenge for LLPs particularly in the subtitled clip, as low comprehension scores illustrate in Table 7.3. In this instance of multimodal irony, voice-over translation resulted in a better understanding of the intended meaning by the participants with low level of proficiency in English.

Table 7.3 Mean comprehension scores of HLPs, MLPs and LLPs in the subtitled (VO_SUB) and voiced-over (SUB_VO) versions (Appendix 8)

Mean	HLPs	MLPs	LLPs
Subtitles VO_SUB	4.5	3.6	3.1
Voice-over SUB_VO	3.8	4.0	2.0

From Table 7.3 we learn that HLPs and MLPs understood the intended ironic meaning considerably better with subtitles (mean=4.5 and mean=4.0, respectively) than with voiced-over (mean=3.8 and mean=3.6, respectively). Relatively higher comprehension scores in the subtitled version are in line with the eye-tracking data, indicating that the processing of the content of subtitles did not result in a high cognitive load, as showed by the relatively short total fixation duration in the subtitle area (see Subsection 6.2.3). As a result, subtitles dragged HLPs and MLPs' attention automatically when these popped out on the screen in ae medium close-up shot. HLPs and MLPs could also look at them out of curiosity to compare the translation of the source text with the target text. As shown in their responses to Q1 and Q2, HLPs and MLPs (e.g., P02B) provided coherent and correct responses describing Sherlock's attitude of dissociation towards Watson and did not have difficulties to understand the meaning of "szanse" his scene. For instance, in response to Q1, which asked why Sherlock is producing the ironic utterance towards Watson, P02B says: "Bo Watson nie trafił za pierwszym strzałem" (BT: "Because Watson did not hit with the first shot"). In response to Q2, which asked the participants to describe Sherlock's attitude

towards Watson, P22B explains "szyderczy, Sherlock wyśmiewa się z jego umiejętności strzelania" (BT: "scoffing, Sherlock is mocking his shooting skills"). In addition, longer fixation durations on Sherlock and Watson's face in the subtitled and voiced-over clips enabled HLPs and MLPs to perceive simultaneously both acoustic and visual cues that contributed to the recovery of Sherlock's dissociative attitude, as revealed in their responses to Q3. To name just a few, Sherlock's tone and volume of voice, music, sounds, special effects in tandem with his facial expressions, impulsive movements, arrogant behaviour and props (e.g., a pipe) were particularly important for HLPs and MLPs in the subtitled and voiced-over clips (e.g., P22A; P07A; P316A; P11B). For example, P22A pointed out "mimika twarzy Sherlocka, napięcie mięśni i żywiołowe ruchy głową. A także wysoki ton wypowiedzi" (BT: "Sherlock's mimic, muscle tension and vivid head movements as well as high tone of voice"), while P11B mentioned "intonacja, wyraz twarzy Holmesa" (BT: "intonation, Holmes' facial expression"). Interestingly, only 1 HLPs out of 44 participants (e.g., P33B) in both groups indirectly acknowledged the importance of editing and camera techniques in the scene under analysis. This view is supported by P33B, who noted: "Szybkosc zmian scen podczas walki; szybkie operowanie kamery" (BT: "Speed of shot changes during a fight; fast camera movements") confirming that viewers mostly fail to perceive editing changes (Szarkowska, Kruger and Krejtz, 2015). The possible explanation for lower comprehension scores in the voiced-over version can be found in responses to Q4. A number of HLPs and MLPs (e.g., P15A; P01A; P08A) reported that they encountered difficulties to retrieve ironic meaning conveyed through voice-over because of the narrator's voice that hindered the reception of the original sounds and dialogues, and thereby made it more difficult to perceive the on-screen characters' tone of voice and emotions.

On the other hand, LLPs found it more challenging to grasp the ironic meaning in the subtitled clips (mean=2.0) than in their voiced-over counterparts (mean=3.1). Although LLPs' fixation values are similar to the ones generated by HLPs and MLPs, in this case the processing of subtitles imposed an additional cognitive load on LLPs, as reflected in their responses to Q1 and Q2. For example, in regard to Q1, most of LLPs were not able to explain the circumstances that lead Sherlock to express his ironic statement towards Watson, while in Q2 LLPs misinterpreted Sherlock's intentions in the subtitled clip and described it literally or opted for the "I don't know" response (e.g., P18B, P22B, P30A). In the voiced-over clip, however, LLPs largely depicted Sherlock's attitude precisely (e.g., P21A; P20A; P03A). LLPs' lower scores in irony comprehension can be explained by the fact that very few participants spotted some non-verbal visual semiotic resources in the subtitled clip under

analysis. That is, only 3 out of 6 individuals noticed Sherlock's gestures, behaviour or the tone of his voice in the subtitled version (i.e., P13B; P18B; P25B). The remaining 4 participants did not mention any non-verbal resources that could help them infer that Sherlock's statement is intended as ironic (e.g., P08B; P21B; P22B). In the voiced-over version, however, most of LLPs accessed multimodal irony through Sherlock's body movement, facial expressions or settings (e.g., P14A; P26A; P30A). In addition, 4 out of 8 individuals also retrieved Sherlock's tone of voice (e.g., P20A; P26A; P03A), despite the voice-over narrator partially hiding the original soundtrack.

7.3 Reception of Acoustically Construed Multimodal Irony

The current section sets out to analyse the retrieval of ironic meaning in these clips in which the on-screen characters' attitude of dissociation is principally relayed acoustically, and where the visual cues are considered a secondary source of information. It then describes in greater detail the process of irony comprehension in the subtitled and voiced-over excerpts of SH1 and SH2 (i.e., SH1_1; SH1_3; SH2_2). Finally, it gauges the extent to which the respondents' level of English proficiency and the modality of audiovisual translation used in each clip influence the interpretation of multimodal irony and the participants' perception of non-verbal resources at play.

7.3.1 Reception of clip: Cemetery and Lord Blackwood

Overall, **SH1_1** features an instance of multimodal irony in which the overwhelming majority of the participants, particularly HLPs and MLPs, decoded the intended meaning, i.e., Sherlock's dissociative attitude to the proposition echoed towards Lestrade, as the questionnaire data presented in Table 7.4 illustrate. My predictions regarding LLPs in the SUB_VO group, however, have been only partially fulfilled. That is to say, very few subjects with low level of English provided accurate responses in the subtitled version; consequently, many of them failed to identify the filmmakers' intentions in this scene. In the VO_SUB group, on the other hand, LLPs provided substantial evidence that they comprehended the intended meaning in the voiced-over version achieving the same high level of understanding as HLPs and MLPs, as presented in Table 7.4.

Table 7.4 Mean comprehension scores of HLPs, MLPs and LLPs in the subtitled (SUB_VO) and voiced-over (VO_SUB) versions (Appendix 8)

Mean	HLPs	MLPs	LLPs
Subtitles SUB_VO	4.5	3.9	2.1
Voice-over VO_SUB	3.8	4.3	4

What stands out in Table 7.4 is that HLPs and MLPs' capacity to retrieve multimodal irony was very similar when watching the clip with subtitles (mean=4.5 and mean=3.9, respectively) and voice-over (mean=3.8 and mean=4.3, respectively). As outlined in my analysis of the eye-tracking data (Subsection 6.3.1), HLPs scanned – rather than read – the subtitles and, as a result, a higher cognitive effort was not required for them to achieve optimal comprehension. MLPs invested slightly more time and effort in understanding the content of subtitles, which might have had an impact on their scores, as demonstrated in Table 7.4. HLPs and MLPs' understanding of the echoic nature of the emotional engagement between Sherlock and Lestrade is also visible in their responses to Q1 and Q2 (e.g., P01A; P07A; P22A; P05B; P06B; P24B). For example, in regard to Q1, P06B explained why Sherlock is expressing his attitude of dissociation to the proposition echoed to mock Lestrade because "Ponieważ policjanci stoją wmiejscu i nic nie robią" (BT: "Because the policemen just stand in the place and do nothing"). Then in Q2 P06B describes him as "Protekcjonalny, pełen ironii i powątpiewań w umiejętności policji" (BT: "Protective, full of irony and doubts in the policemen's skills"). HLPs and MLPs (e.g., P24B; P05A; P03B; P06B) also found visual information in the centre of the screen highly engaging in the subtitled and voiced-over clips, as the fixation values suggest (see Subsection 6.3.1). Hence, the participants supported their understanding of irony with a wide array of non-verbal semiotic resources pertaining to mise-en-scène and cinematography, as reported in Q3. For example, HLPs and MLPs mentioned gestures, eye contact, head movement or smiling, as illustrated by P24B: "Delikatny uśmiech, nieco uniesiona głowa i oceniające spojrzenia" (BT: "A delicate smile, slightly raised head and judgemental looks") or P05A: "Gesty i ruchy ciała - pozycja Holmes'a w rozmowie z Lestradem, stał bokiem. Ujęcie kamery - to na Lestrada, policjantów, to na grób, co wskazywało tylko na ich niemoc" (BT: "Gestures and body movements - Holmes' position in conversation with Lestrade, standing aside. Changes of the camera shots - first it was directed to Lestrade, to the police and to the grave in order to show their idleness"). What is more, both HLPs and MLPs acknowledged the significance

of the acoustic modes in this scene, which predominantly construed Sherlock's dissociative attitude towards Lestrade. Particularly, they noticed Sherlock's tone and intonation of voice, as mentioned by P05A: "Dźwięk - Ton głosu Sherlocka w rozmowie z Lestradem, niepochlebny, ironiczny" (BT: "Sound - Sherlock's tone of voice in conversation with Lestrade, unflattering, ironic"). As anticipated, more subtitling than voice-over viewers grasped the characteristics of Sherlock's voice modulation, even though the discrepancies between both groups were not that major (P01A; P05A; P06A; P22A in SUB_VO; P06B; P10B; P07B in VO_SUB).

On the other hand, LLPs' comprehension of the intended meaning relayed with voice-over was roughly similar to that of HLPs and MLPs (mean=4.0). When viewing the subtitled clip, however, LLPs provided far more false or imprecise responses to the questions probing irony comprehension (mean=2.1). As the participants were processing the subtitled text much slower than HLPs (see Subsection 6.3.1), they had to elaborate more extensively on the subtitles; consequently, processing the intended meaning required a bigger cognitive effort. As a result, only 3 out of 8 LLPs were able to explain the circumstances that contributed to the construal of multimodal irony in the subtitled clip in their responses to Q1 and Q2 (e.g., P21A; P26A; P30A). Other LLPs gave either false or inconclusive responses, e.g., P30A: "Jak wynieść trumnę żeby nikt nie widział" (BT: "How to get the coffin out of the grave so that no one could see it") or P21A: "Znają się dobrze, że nawet prywatnie, ale nie są w stosunku do siebie zbyt wylewni. Koncentrują się na śledztwie" (BT: "They [Sherlock and Lestrade] know each other very well and are not effusive to one another even in privately. They concentrate on the investigation"). In the voiced-over clip, 5 out of 6 LLPs provided evidence that they accessed the intended meaning, describing Sherlock's attitude of dissociation as ironic (e.g., P14A; P08B; P13B), and thus earning more comprehension points. As far as Q3 is concerned, LLPs identified a great number of non-verbal visual resources, both in the subtitled and voiced-over clips. Among others, several viewers pointed out Sherlock's gestures, body movements (e.g., a straight posture, nonchalant behaviour), mimics, eye contact (i.e., looking down at Lestrade and looking from sunglasses) or physical distance between Sherlock and Lestrade (e.g., P05A, P04A, P01B, P10B), despite the time spent on reading the subtitles. When it comes to the leading acoustic mode in this scene, only some LLPs extracted information conveyed through that channel. To be more precise, 4 out of 8 LLPs (P03A; P13A; P20A; P26A) spotted the tone of voice used by Sherlock in the subtitled clip, in contrast to only 2 out of 6 LLPs in the voiced-over clip (P13B; P17B). Other LLPs admitted in their responses to Q4 that they struggled to grasp features of voice

modulation because of the voice-over narrator hinders the reception of the original dialogue. For example, P13B revealed that "utrudniło mi to usłyszenie tonu głosu i spowodowało dekoncentrację poprzez dwukrotne słyszenie tych samych treści" (BT: "It made it more difficult for me to hear the tone of voice and distracted my attention as I heard the dialogue twice"), while P18B commented "lektor zagłusza mowę aktorów" (BT: "The narrator drowns the actors' voices").

7.3.2 Reception of clip: Sherlock and Lord Coward

Sherlock's disapproving attitude in clip **SH1_3** was on average interpreted by all viewers as ironic both in the subtitled and voiced-over version, as predicted. Table 7.5 demonstrates their successful understanding of irony across all subgroups.

Table 7.5 Mean comprehension scores of HLPs, MLPs and LLPs in the subtitled (SUB_VO) and voiced-over (VO_SUB) versions (Appendix 8)

Mean	HLPs	MLPs	LLPs
Subtitles SUB_VO	4.7	4.2	4.1
Voice-over VO_SUB	4.6	4.8	4

The data in Table 7.5 clearly show that the overwhelming majority of the target audience retrieved the intended ironic meaning in the subtitled version successfully, despite HLPs, MLPs, and in particular LLPs' longer total and mean fixation durations in the subtitles area. Thus, in this clip, many participants did not fixate on the subtitles because of their content but because of their location on the screen, I argue. That is to say, the viewers tend to read the subtitles, when they are positioned right under the on-screen character's face in a medium close-up or close-up shot, as is the case in SH1_6, although they can be also straightforward to understand. Precise and complete responses to Q1 and Q2 also confirm that the deletion of the question tag "Is there?" in the subtitles did not affect comprehension of multimodal irony, contrary to my assumptions. For instance, in Q1, many of those surveyed correctly explained why Sherlock acted this way towards Lord Coward (e.g., P01A; P19B; P24B; P19A; P04A; P25A), like P25B: "Bo zostało niewiele czasu na rozwiązanie zagadki" (BT: "Because there was little time left to solve the puzzle"). In Q2, HLPs, MLPs and LLPs accurately described Sherlock's attitude towards Lord Coward as disrespectful, mocking and ironic (e.g., P25B; P18B; P11B; P24B; P08A; P09A; P01A) as P25B did: "Lekceważy go

uśmiechając się ironicznie. Traktuje z wyższością" (BT: "[Sherlock] is disrespecting him by smiling ironically. He treats him with a sense of superiority").

Nevertheless, I have observed some differences between HLPs, MLPs and LLPs in the perception of non-verbal resources in the subtitled version and in its voiced-over counterpart. Q3 shows that 7 out of 15 HLPs and MLPs in SUB_VO recognised the music score and Sherlock's tone of voice, which were considered vital to the discovery of multimodal irony in SH1_1 (e.g., P01A; P09A; P15A; P03A). In contrast, in the VO_SUB group, only 1 out of 8 HLPs (P01B) found the film score helpful to retrieve multimodal irony, while 3 other HLPs and MLPs referred to Sherlock's tone of voice only (e.g., P05B; P10B; P17B), due to the voice-over narration that reduced the audibility of the original soundtrack. Longer fixation durations on Sherlock's face and body (see Subsection 6.3.2) also allowed several participants to spot a number of visual cues such as mimics (e.g., smirking), gestures and body movements (e.g., relaxed and nonchalant posture) both in the subtitled and voiced-over clips. Other HLPs, MLPs and LLPs also noticed the characteristics of settings and objects like a pipe (e.g., P03A; P08A; P01B; P32B; P06A). More importantly, HLPs (e.g., P19A, P22A) described Sherlock's facial expressions in ample detail distinguishing wide-open eyes, "a blank face" and raised eyebrows, which have been identified as critical in the production and reception of irony in the previous studies (Attardo et al., 2003). Hence, many respondents with high level of English noticed specific features of facial expressions that the participants with medium and low level of English totally missed. Only a few individuals reported problems to identify Sherlock's ironic attitude in Q4. In the subtitled clip, 5 out of 21 participants admitted that they could not concentrate on Sherlock's facial expressions and behaviour, as they spent too much time reading the subtitles (e.g., P04A; P30A; P26A in A2.1). In the voiced-over version of the clip, only 3 out of 3 participants complained that the voice-over narrator hampered the reception of Sherlock's voice modulation in the original soundtrack (e.g., P18B, P27B).

7.3.3 Reception of clip: Watson's stag party

In clip **SH2_2**, on the whole, the vast majority of the target audience rightly understood Mycroft's attitude as dissociative, as revealed in the analysis of questionnaire responses to Q1 and Q2. Particularly, HLPs and MLPs earned twice as many comprehension points as LLPs watching the scene with subtitles and voice-over. As anticipated, several subjects with low level of English experienced several difficulties to grasp and interpret the non-verbal acoustic semiotic resources in the voiced-over clip and the non-verbal visual semiotic

resources in the subtitled clip, which affected their successful comprehension of irony, as indicated in Table 7.6.

Table 7.6 Mean comprehension scores of HLPs, MLPs and LLPs in the subtitled (VO_SUB) and voiced-over (SUB_VO) versions (Appendix 8)

Mean	HLPs	MLPs	LLPs
Subtitles VO_SUB	4.5	4.1	2.2
Voice-over SUB_VO	4.1	4	2.5

It is apparent from Table 7.6 that both HLPs and MLPs derived the intended meaning successfully in the subtitled (mean=4.1 and mean=4.0, respectively) and voiced-over (mean=4.5; mean=4.1, respectively) excerpts under scrutiny. In their responses to Q1 and Q2, the majority of HLPs and MLPs explained precisely why Mycroft behaved that way towards Watson and described in detail his attitude as ironic and mocking (e.g., P01A; P07A; P10A; P22A; P06B; P11B; P27B). For example, P01A said: "Wiedział że Sherlock zapomniał o wieczorze kawalerskim" (BT: "He knew that Sherlock forgot about his bachelor party") in reference to Q1 and "Niby uprzejmy ale ironiczny" (BT: "Kind of polite but ironic") in regard to Q2. Thus, HLPs and MLPs' understanding of irony was facilitated by their perception of the acoustic semiotic resources that are highly significant to the narrative in this scene. Thus, several HLPs and MLPs distinguished Mycroft's tone of voice and "antiphonal laughter" construing Mycroft's attitude of dissociation in the subtitled and voiced-over clips (e.g., P05A, P16A, P19A, P09A or P22A). What is more, HLPs and MLPs spent a considerable amount of time fixating on the on-screen characters' faces in the middle of the screen. This gaze pattern enabled HLPs and MLPs to extract information conveyed visually, including features of Mycroft's facial expressions, such as a sneering smile and smirk (e.g., P22A, P03A, P14A, P17B) in the voiced-over and subtitled versions.

Counterintuitively, LLPs understood meaning featuring the use of multimodal irony much less accurately than HLPs and MLPs in the subtitled (mean=2.5) and voiced-over (mean=2.2) clips, although LLPs focused on the AOIs in similar proportions to HLPs and MLPs. As outlined in the questionnaire responses, in respect to Q1, only 3 out of 8 LLPs (P18B; P21B; P25B) in the subtitled clip and 2 out of 6 LLPs (P14A; P21A) in the voiced-over excerpt explained the reasons why Mycroft directed his mocking towards Watson. Some participants demonstrated partial irony retrieval in Q2 alone, describing Mycroft's

attitude accurately as arrogant and mocking (e.g., P03A, P30A). For example, in reference to Q1, P25B answered: "Wie, że Sherlock nie zaprosił jego znajomych. Żartuje sobie z niego" (BT: "He knows that Sherlock did not invite his colleagues. He is making fun of him") and described Mycroft's attitude as "drwiący, żartuje sobie z niego ironicznie" (BT: "mocking, he is joking about him ironically"). Other LLPs misinterpreted Mycroft's attitude, defining it as "przyjacielski" (BT: "friendly") or "współczujący" (BT: "compassionate") (e.g., P08B, P13B, P20A) in both groups. LLPs' failure to retrieve meaning in this scene can thus be linked to LLPs' lower linguistic ability; in other words, their low level of English and the modality of audiovisual translation used in each case made it more difficult for them to grasp the non-verbal acoustic cues in the original dialogue, which are of high narrative saliency in this clip. As illustrated in Q3, none of LLPs identified "antiphonal laughter" or Mycroft's voice qualities in the voiced-over version, while only 2 out of 8 LLPs noticed these acoustic cues in its subtitled counterpart, when the original dialogue was not hidden by the voiced-over narration.

As shown above, LLPs largely failed to grasp information relayed visually and/or acoustically, in contrast to HLPs and MLPs. As a result, LLPs faced more interpretative challenges than HLPs and MLPs to recognise Mycroft's attitude of dissociation and interpret his comment as ironic. It seems that their low level of English of LLPs limited their access to the non-verbal semiotic resources, and thus lowered their abilities to understand a complex phenomenon like irony. What is more, voice-over additionally hampered their perception of acoustic cues and subtitles hindered their perception of visual cues that HLPs and MLPs could easily spot in both versions. Thus, it appears that "the interpretation of emotions might change due to exposure to a foreign language" (Ożańska-Ponikwia, 2013: 8). The participants with high and medium level of English, who are exposed to a foreign language and culture on a more regular basis, can be more likely to understand and associate emotions with a certain type of behaviours or attitudes expressed in the SL and relayed in the TL either with subtitles or voice-over.

In addition, a correlation test has been conducted to verify whether there is any correlation between the participants' language levels and the comprehension scores allocated for their responses in the experimental questionnaire. Thus, the Pearson's correlation test has been carried out and the results have been presented in Figure 7.1 below:

Figure 7.1 Pearson's correlation test between the language levels and the comprehension scores for SUB_VO and VO_SUB groups

```
        Pearson's product-moment correlation

data:  sub_vo_comp$Scores and sub_vo_comp$Groups
t = 3.7019, df = 21, p-value = 0.001322
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.2917668 0.8265093
sample estimates:
      cor
0.628394
```

```
        Pearson's product-moment correlation

data:  vo_sub_comp$Scores and vo_sub_comp$Groups
t = 3.3382, df = 19, p-value = 0.003455
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.2390772 0.8235434
sample estimates:
      cor
0.6080131
```

As revealed in Figure 7.1 above, a strong positive correlation has been found between comprehension scores and language groups ($t(21) = 3.702, p < .05, r = 0.63$) for the SUB_VO group and the VO_SUB group ($t(19) = 3.34, p = < .05, r = 0.61$), as predicted. It means that the level of proficiency in English does play a significant role in the comprehension of multimodal irony and the better the participants know the foreign language, the better results in the comprehension of the intended meaning they are likely to achieve.

Overall, the analysis of the questionnaire responses given by the target audiences in this experiment strengthens the premise that the kinesic, visual and acoustic modes clearly operate in conjunction with film dialogue. These verbal and non-verbal meaning-making resources also intersect to contribute to detection of the on-screen character's ironic intentions and emotions towards their interlocutors in the selected subtitled and voiced-over clips of SH1 and SH2. Nevertheless, some of these non-verbal resources construing the intended ironic meaning visually and/or acoustically have been interpreted differently in various cultural circles, as outlined in previous research (Antonijevic, 2008; Williams, Burns and Harmon, 2009). In the next section, these differences in the retrieval of certain non-verbal elements are discussed and, by extension, the emotions associated with them in the construal of irony between the Polish and American audiences, drawing on the reception of the non-verbal markers of irony by the American viewers analysed in an earlier study by Attardo et al. (2003).

7.4 Reception of Multimodal Irony across Cultures

In this section, it is debated whether the interpretation of non-verbal resources construing multimodal irony on screen differs across cultures. Specifically, the aim of this section is to conduct a comparison to see how viewers in the current experiment grasped multimodal irony with non-verbal information available in the selected clips of SH1 and SH2 and contrast it with previously existing research.

Past studies have predominantly touched on the role of acoustic cues (e.g., voice modulations) in the construal of ironic utterances in languages like English, German or Italian (Rockwell 2000; Bryant and fox Tree 2005; Scharrer and Christmann, 2011; Antolli et al., 2000). It has been decided, instead, to contribute to the limited body of research on non-verbal resources that are used to construe irony visually. The focus has been explicitly set on the importance of facial expressions for the detection of the on-screen characters' dissociative attitude since these semiotic resources are considered essential in the construal of irony (e.g., Attardo et al., 2003). As no study has been undertaken so far, to the best of my knowledge, to measure the reception of irony in SH1 and SH2 by American audiences, the present study has relied on the existing study of visual markers of irony conducted by Attardo et al. (2003). In his analysis, American native speakers were asked to describe specific features of facial expressions of ironical speakers from television situation comedies. In the course of this study, American viewers recognised "a blank face" raised eyebrows and a smile, among others, as markers of ironic meaning on screen.

The post-experimental questionnaire responses showed (Appendix 8) that the findings are consistent with the data obtained in Attardo et al.'s (2003) analysis, although the subjects were not informed about the aim of the study or about the type of facial emotions that they were exposed to. Nevertheless, similarly to American respondents, several viewers also identified smiles as one of most meaningful non-verbal elements in the reception of multimodal irony. Many of them also distinguished the type of smile they observed, namely, ironic, mocking, sneering or smirking (e.g., P04A; P18A; P07A; P14 A; P24B; P10B; P16B in SH1_2; P034A in SH2_1 or P04A; P08A; P22A; P30A, P14A, P03A in SH2_2). Another non-verbal semiotic resource that influenced the ease of irony comprehension in Polish and American viewers alike are raised eyebrows (e.g., P22B in SH1_2; P19B in SH2_1), although only a small number of Polish viewers was able to retrieve this feature of facial movements. The presence of subtitles on the screen could hinder their perception of more specific details of face expressions in the retrieval of the intended meaning. What is more, all respondents also acknowledged the significance of "a blank face" in the reception of

multimodal irony. Although this type of emotionless face has been largely unexplored in the field of pragmatics or linguistics so far, the evidence derived from this study verifies Attardo et al.'s (2003) study to some extent. That is to say, 1 out of 44 participants (P22A in SH1_3) recognised Sherlock's dissociative attitude through his emotionless face that involves "no smile, no grimace, no eyebrow raising, no frown" (Attardo et al., 2003: 254) and labelled it with a Polish equivalent, that is, "a stone face", known as "emotionless and stillness" (online dictionary PWN). This finding suggests that both Polish and American cultures have a common understanding of the notion of a blank face in the retrieval of multimodal irony.

Let us now consider more meticulous responses collected from the participants in Q5 in the present study. In this question, all viewers were provided with a full list of non-verbal resources pertaining to mise-en-scène, cinematography, sound and editing that smooth discovery of ironic meaning. In Table 7.7, I have presented only those non-verbal elements that refer to the peculiarities of the on-screen character's facial features drawing on Attardo et al.'s (2003) study.

Table 7.7 The percentage breakdown of non-verbal resources pertaining to the on-screen characters' facial expressions that 44 participants spotted in the experiment

Participants	Kinesic semiotic resources in the percent ratio							
HLPs **MLPs** **LLPs**	Blank face	Nodding	Smiling/Smirking	Winking	Rolling eyes	Raised eyebrows	Lowered eye brows	Wide-open eyes
	45%	16%	84%	9%	36%	57%	30%	36%

It can be seen from the data in Table 7.7 that many subjects reported the same features of facial expressions that they mentioned in the open-ended Q3. One anticipated finding is that 84% of the viewers identified a smile/a smirk as a most salient non-verbal resource in the reception of irony on screen. The second most significant non-verbal element mentioned by 57% of HLPs, MLPs and LLPs were raised eyebrows, while 45% of those surveyed identified "a blank face" as a significant resource that contributes to detection of ironic meaning, which is somewhat surprising, as in Q3 only one participant noticed it. The questionnaire data presented in Table 7.7 also indicates that the participants discerned a great number of other non-verbal resources in irony comprehension. For example, 36% of those surveyed singled out wide-opened and rolling eyes, while 30% of the respondents considered lowered eyebrows a vital visual marker of ironic meaning. Finally, nodding and winking were found by many viewers as the least important (16% and 9%, respectively).

These results further support the idea that both Polish and American audiences are equipped with a range of non-verbal semiotic resources like smile/smirk that relay a similar set of emotions in both cultures. Although both audiences come from remote countries and

speak typologically very different languages, their interpretation of smile/smirk in the context of multimodal irony is very similar and can be included in "the same basic emotional repertoire" (Evans, 2001: 11). Interestingly enough, even specific non-verbal elements were also interpreted in a similar way across the two languages. Although the concept of "a blank face" has been found to signal the on-screen character's ironic statement by the American viewers, the Polish ones identified the equivalent term, namely, a stone face and this "overlap of concepts in L1 and L2 may facilitate positive transfer in the L2 learning process" (Ożańska-Ponikwia, 2013: 9).

Hence, these non-verbal cues as well as others like wide-opened eyes or raised eyebrows were identified in the relay of the attitude of dissociation by both audiences. Although the notion of irony is frequently described as a culture-specific phenomenon (e.g., Barbe, 1995), the findings of this comparative study suggest that the two culturally very different groups of viewers share some of their understanding of multimodal irony through a common base of non-verbal features of facial expressions.

7.5 Conclusions

The main goal of Chapter 7 was to gauge to what extent the target audience recognised the attitude of dissociation, and thus retrieved the intended multimodal irony in the subtitled and voiced-over clips of SH1 and SH2. It has been further analysed what non-verbal resources contributed to the identification of ironic attitude by the participants with high, medium and low degree of proficiency in the English language and to what extent they were able to grasp them when ironic meaning was relayed with subtitles and voice-over. The analysis has also been complemented with the eye-tracking data to verify the assumptions that I have made in the previous chapter.

The results from the analysis of questionnaire data have crucial implications for the retrieval of irony and have confirmed that "irony comprehension is a complex cognitive task" (Hala et al., 2010: 299). The most interesting empirical finding is that the modality of translation and the level of English proficiency have been found to have an impact on the comprehension of multimodal irony in the clips under analysis. On average, HLPs and MLPs achieved relatively high irony comprehension scores both in the subtitled and voiced-over clips, as expected. In some instances of multimodal irony, HLPs and MLPs earned even more comprehension points when watching the clips with subtitles than voice-over, although differences are subtle. As HLPs and MLPs predominantly fixated on the AOIs in the centre of the screen, they managed to extract information relayed visually or acoustically rather

effortlessly even in the presence of subtitles and voice-over in the clip, as presented in this chapter. Thus, they probably either looked at subtitles at times to reconfirm the words they misheard in the original soundtrack or they simply read the subtitles automatically, as past studies have shown (e.g., Orrego-Carmona, 2015). This viewing behaviour allowed them to perceive a great number of visual non-verbal resources in the subtitled and voiced-over version. In addition, HLPs were able to provide more specific features of facial expressions such as "a blank face" and raised eyebrows or describe body language in abundant detail, which also seems to contribute to their comprehension scores. Some of HLPs and MLPs also handled ironic meaning relayed acoustically in the voiced-over version, that is, the on-screen characters' voice modulation (e.g., tone of voice) and the characteristic quirky film score that highlighted the ironic nature of Sherlock's dialogue acoustically. Because of their exposure to a foreign language and culture, HLPs and MLPs appear to be more likely to identify non-verbal cues loaded with culture-specific emotions and meanings, facilitating irony interpretation in a non-native language irrespective of whether multimodal irony is relayed through subtitles or voice-over. Thus, English proficiency along with socio-cultural and sociolinguistic competence "not only broaden knowledge of cultural and social norms present in every society but also enlarge the emotional repertoire by means of which acquisition of new concepts takes place", ultimately paving the way for successful irony comprehension (Ożańska-Ponikwia, 2013: 6).

While HLPs and MLPs understood the ironic meaning to a similar extent, some vital differences have been detected in the retrieval of multimodal irony among LLPs. Overall, LLPs provided far less accurate responses than HLPs and MLPs, particularly in the subtitled clips (i.e., SH2_1; SH2_3; SH1_1). A possible explanation for these results is that subtitles are believed to place an extra effort on the system of visual perception and impair LLPs' cognitive abilities to process on-screen information (Szarkowska et al., 2014; Smith, 2014). As a result, "reading subtitles while watching a film is a more demanding activity than watching dubbed films" or in the context of this study voiced-over films (Orrego-Carmona, 2015: 244). Nevertheless, the deficits in irony comprehension have also been found in other instances of multimodal irony in the subtitled and voiced-over clips. LLPs noticed some general non-verbal markers such as mimics, gestures or body movements that were expected to facilitate detection of irony. In some cases, however, they did not interpret the intended meaning of these non-verbal cues, for instance, when contrasting Watson's ironic and Sherlock's looks in SH1_2, to come to the conclusion that the intended meaning should not be understood as literal. It is argued that some of these linguistic misunderstandings are due

to the fact that LLPs struggle to acquire new expressions of facial emotions, for instance, that are key to the comprehension of multimodal irony. Thus, it seems that low level of English proficiency impedes the understanding of irony.

Last but not least, a comparison study of non-verbal resources in the reception of multimodal irony has revealed that Polish and American native speakers share a similar understanding of some most meaningful features of facial emotions underpinning irony interpretation. That is to say, "a blank face", raised eyebrows, wide opened eyes or a smile have been found to facilitate the retrieval of multimodal irony in both audiences. These findings raise intriguing questions regarding the nature of multimodal irony as a culture-specific phenomenon construed by a set of recognised and recognisable non-verbal elements, which could be addressed in future research.

The results derived from the analysis of questionnaire responses in this chapter that complemented the examination of the eye tracking data in Chapter 6 and the multimodal data in Chapter 5 have provided me now with a comprehensive overview of the construal and reception of multimodal irony by viewers in film translation. In the next chapter, the research questions have been answered on the basis of the examined experimental data.

8 CONCLUSIONS

The overarching aim of this research project was to gauge the extent to which verbal and non-verbal resources contribute to the comprehension of multimodal irony in selected excerpts of subtitled and voiced-over films. Another goal was to explore the impact of English proficiency on irony comprehension when watching the subtitled and voiced-over clips under scrutiny using a triangulation of research methods. Methodological triangulation allowed me to examine the construal, relay and retrieval of 6 instances of multimodal irony in my data set consisting of SH1 and SH2. Drawing on insights from cognitive pragmatics, multimodality and film studies, the present thesis constitutes a truly interdisciplinary contribution to scholarly research of multimodal irony. The findings summarised in the following sections provide an answer to each of my research questions; each of my secondary research questions will be addressed before answering the central one.

8.1. Findings of this Study

In order to answer the overarching research question – i.e., What is the contribution of verbal and non-verbal resources to irony comprehension by Polish viewers in the subtitled and voiced-over Polish versions of *Sherlock Holmes* (2009) *and Sherlock Holmes: A Game of Shadows* (2011)? – it has been endeavoured to respond the secondary questions:

1.　　How is irony construed in the two Sherlock Holmes films?
　　　a) How do verbal resources contribute to the construal of irony?
　　　b) How do non-verbal resources contribute to the construal of irony?

The first and most obvious finding to emerge from the multimodal analysis of the descriptive data set is that irony is construed multimodally by the filmmakers in the Sherlock Holmes films. Film dialogue clearly plays a pivotal role in the generation of irony and foregrounds linguistic indirectness that operates in the two layers of film communication, i.e., between the on-screen characters in the film text and between the on-screen characters and viewers of audiovisual productions. Like in other manifestations of linguistic indirectness, e.g., implicature (Desilla, 2009), irony has also been found to assist with the creation of humour in the two films under scrutiny. As shown in this thesis, the screenwriters have a plethora of linguistic tools at their disposal, such as ironic repartee or wordplay, with particular emphasis on banter and cut wit that trigger a comic effect when expressing an attitude of dissociation – for instance, to mock the appearance, competence or intentions of another on-screen

character. The presence of "the contest of wit" (Abrams, 1993: 220) is mainly demonstrated in the type of the relationship between Downey's Holmes and Law's Watson, which signals the special ironic humorous character of their 'bromance'. This close relation between irony and humour has also been illustrated through the use of incongruity and superiority (Vandaele, 1999/2002). The screenwriters thus used ironic metaphor and ironic ripostes in the film dialogue to create humorous irony so that viewers grasp the incongruity between literal and non-literal meaning which contradicts their expectations. In the double-layer of film communication, humorous irony is also construed through superiority, where a character understands the literal meaning while the audience detects the speaker's attitude of dissociation. This is clearly illustrated by the dialogue exchanges between Sherlock and Lestrade, during which the inspector does not realise that he is being mocked. Sarcasm is also of particular importance in the composition of humour in the two Sherlock Holmes films that screenwriters employed to amuse the audience. This is evident, for instance, in Sherlock's mockery of Lestrade through a range of satirical or ironic remarks. As irony is particularly celebrated in romantic comedies (Kozloff, 2000), multimodal irony is predominantly observed in action comedies, where it is composed through a whole range of non-verbal semiotic resources. As illustrated in this thesis, the film dialogue is not capable of relaying ironic meaning in isolation from the visual, kinesic and acoustic modes.

On the whole, it became evident that multimodal irony fulfills comedic and narrative functions always in concert with mise-en-scène, cinematography, editing, and/or non-verbal soundtrack, as expected. With regard to the visual modes, the multimodal analysis revealed that certain camera functions – one of the manifestations of the cinematography mode – such as camera angle and distance are particularly significant in the production of ironic meaning on screen (Dix, 2008). For instance, the director of both films used a low camera angle in the construal of Sherlock's dissociative attitude, strengthening his superiority and dominance over other protagonists; in other cases, he used a high camera angle to make the on-screen characters look less powerful towards their interlocutors. This use of the cinematographic resources is particularly visible in the conversations between Sherlock, Watson and Lestrade.

Camera distance has also been found to be very useful in the composition of ironic meaning on screen. For example, medium close-up and close-up shots gave prominence to specific features of facial expressions including eyes, gaze, eyebrows, mouth or head movements, which have been considered vital to the recognition of the on-screen characters' emotions that are used to express an attitude of dissociation. In long and medium long shots,

196

on the other hand, the viewers could grasp body movements, gestures, gaze direction between the on-screen characters and elements that comprise the filming scene such as the settings or the props like garments used to construe the detective's dissociative attitude. The feeling of tension, suspense and excitement enhanced by frequent and abrupt editing shots have also come across as ancillary in the relay of ironic meaning on screen.

As far as the visual modes are concerned, several non-verbal semiotic resources pertaining to mise-en-scène have also been seen to compose the intended ironic meaning, namely, the settings and the props (like pieces of clothes, a pipe), as mentioned above. In the present study, the kinesic modes have also been considered highly productive in the expression of the attitude of dissociation, with particular emphasis on the on-screen character's facial expressions and body language. For example, the non-verbal semiotic resources of facial expressions identified by Attardo et al. (2003) such as "a blank face", raised eyebrows, widely open eyes and smiling or smirking have also been found to play a role in my own descriptive data analysis.

The acoustic modes have also signaled the presence of the intended ironic meaning on several occasions, primarily by intensifying emotions, feelings and attitudes of the protagonists through non-verbal elements of speech, music and special effects. In the multimodal analysis, it has been confirmed that specific features of voice modulation like flat and raising intonation, heavy stress, medium pitch, slow speech rate and pauses are highly significant in the expression of the dissociative attitude in the film text under study, as had been found to be the case also in previous studies. The presence of multimodal irony is additionally highlighted by "antiphonal laughter" (Bryant, 2011) that has also been found in the descriptive data set. The dissection of descriptive data also brought to the fore instances of multimodal irony where music was of paramount importance. The results obtained from the multimodal analysis of the intended ironic meaning are thus consistent with those of Attardo et al. (2003), Boggs and Petrie (2008) or Turner (2015), to name just a few, who identified a wide array of non-verbal semiotic resources of the language of film that contribute to the production of ironic meaning on screen.

2. How is irony relayed in the Polish subtitled and voiced-over version of the two Sherlock Holmes films?

The second set of findings refers to the way ironic meaning is relayed in the subtitles and voice-over of the two Sherlock Holmes films. The descriptive analysis of the film dialogue

translated from the SL into the TL revealed two broad categories of irony relay in the film texts under study, namely, *preservation* (i.e., the wording and the form of the SL have been preserved in the TL) and *modification* (i.e., the wording and the form of the SL have been modified in the TL) in the subtitled and voiced-over clips extracted from SH1 and SH2. In regard to the subtitled clips, it has been observed that preservation was sometimes opted for by the subtitlers; specifically, in 2 out of 6 clips each word of the SL dialogue rendered into the TL subtitles. That is to say, occasionally, verbatim subtitles from the SL into the TL were produced, when this was allowed by time and space limitations or when it was certain that the intended meaning expressed in the SL will be equally understood in the TL. In this case, the subtitler did not have an opportunity to manipulate the intended ironic meaning that the filmmakers had intended to convey. The pattern of preservation suggests that the original dialogue that was translated literally should also retain the ironic effect as it is intended in the film dialogue.

Modification, on the other hand, was chosen in the majority of cases, that is in 4 out of 6 clips, which confirmed myexpectations. This finding is in line with a number of previous studies confirming that the reduction, paraphrase, deletion or condensation of the original dialogue to interlingual subtitles is a well-known and widely accepted translation strategy (see Georgakopoulou, 2010; Tomaszkiewicz, 2006). More importantly, this result is also consistent with the research of Pelsmaekers and Van Besien (2002) in which they also recognised that modification is the most frequent technique to relay irony in subtitles. In the current study, the changes that have found their way into the TL at textual or sentence levels include paraphrase, condensation or reduction. A possible explanation for the strong trend towards the text modification in subtitles can be linked to the impact of spatio-temporal constraints on the production of subtitles. Considering the limited space at the bottom of the screen and the fast pace of the source dialogues, text reduction seems to be inevitable. In the case of long dialogues, more text reduction is required, otherwise, the subtitlers would run the risk of displaying subtitles on screen for such a short time that viewers would be unable to read them and follow the action in the middle of the screen.

Apart from these more technical aspects, film dialogue is also frequently modified as part of the subtitling process because of idiomatic expressions or culture-specific expressions, as illustrated in the descriptive data set. Since the comprehension of irony itself is a complex task, the subtitler either omitted these words or substituted them with other more colloquial phrases available in the target text. In this way, s/he decreased the level of linguistic complexity and made the subtitled content easier and faster to process. Although,

the subtitler's mediation may be detrimental to the perception of the ironic effect of the TL, vis-à-vis the ironic effect that the film director used in the SL, leaving longer stretches of dialogue untouched in the subtitles could significantly hamper the comprehension of the intended ironic meaning. Nonetheless, preservation and modification can jeopardise the comprehension of ironic meaning in subtitles, if viewers are not able to process a wide range of non-verbal elements that support the transfer of irony visually and/or acoustically.

Surprisingly, a tendency to preserve ironic meaning was noticed in the voiced-over versions of SH1 and SH2. That is, in 3 out of 6 instances, the SL dialogue containing multimodal irony was rendered word-for-word, while the remaining 3 clips were modified in the voiced-over version. This shift can be linked to the conventions about space and time constraints that are not as rigid in voice-over as in subtitles, although voice-over also requires a high degree of condensation or reduction of the source text to facilitate access to the source dialogue (Woźniak, 2012). Since speech typically communicates more information than written language, the voice-over narration delivered more information than the subtitles. For instance, the question tag that was omitted in the subtitled version was kept in its voiced-over counterpart. Similarly, to the subtitles, idiomatic or cultural expressions identified in the original dialogue were either reduced or replaced with high-frequency and more conversational phrases in the voice-over.

3. How do Polish viewers, representing different levels of English proficiency, consume the intended ironic meaning in the subtitled and voiced-over Sherlock Holmes films, as shown by eye-movement data?

The experimental component of the thesis used eye-tracking technology to investigate how viewers watched the action on the screen in the subtitled and voiced-over clips of SH1 and SH2 that feature ironic meaning. Additionally, it also aimed to test the assumptions on which the analyst had conducted the multimodal analysis during the descriptive strand of this project. In the analysis of eye-tracking data, major discrepancies were observed, not only between the subtitled and voiced-over excerpts, but, interestingly, within audience groups with high, medium and low levels of English proficiency. To date, the eye-tracking studies have explored viewing behaviour of participants with high and low level of English-language competence (e.g., Orrego-Carmona, 2015), while participants with medium level have been ignored so far. As shown in the present study, there is a substantial proportion of all viewers that have a good command of English at B1 or B2 level (medium) which

supported the inclusion of this language subgroup in the current experiment. The most obvious finding to emerge from the analysis of eye-tracking data refers to the differences in the viewing behaviours when watching scenes with subtitles and voice-over.

When watching subtitled clips, the participants overall actively explored the scene, dividing their attention between reading the subtitles and following the action in the middle of the screen to various extents depending on their language skills. In general, the viewers with high level of English fixated more and longer in the centre of the screen, as predicted, looking at the AOIs around the character's faces the most frequently. As a result, they were more likely to grasp non-verbal features of facial expressions in ample detail, which contributed to a better comprehension of irony on their part. On the whole, HLPs spent much less time gazing at subtitles than their counterparts with lower linguistics abilities. As a result, HLPs looked at the subtitle area on specific occasions only. These included, for example, these instances in which they could extract or re-confirm information they misheard or wanted to compare in the original soundtrack. Occasionally, HLPs fixated on the subtitles longer than expected, simply by appearing on screen. As a result, HLPs could adapt their subtitle-reading performance to their needs. As predicted, HLPs spent less time reading the subtitles but more time exploring the AOIs in the middle of the screen, as the eye-tracking data indicates, although theses fixation values were not found to have statistical significance.

The experimental data also revealed that MLPs exhibited varied gaze patterns across all 6 clips, particularly in respect to the subtitled videos. On the one hand, MLPs behaved in a manner similar to HLPs in the majority of the instances and avoided dwelling on the subtitles for a substantial amount of time. On the other hand, however, MLPs fixated on the subtitle area nearly as long as LLPs, particularly in the case of subtitles where the content of the original dialogue was modified suggesting "more effortful cognitive processing" (Holmqvist et al., 2011: 381). Since MLPs also distributed their attention more evenly on the screen than LLPs, they could spot some of the non-verbal semiotic resources when they explored the AOIs in the middle of the screen. Thus, my assumptions were confirmed only partially, as the fixation durations of MLPs were statistically significant in contrast to HLPs and LLPs.

Participants with low level of English, on the other hand, mostly focused their attention on the subtitle area and made considerably longer and more frequent fixations when processing the subtitled text. These gaze patterns may be associated with reading difficulties and indicate that the processing of the subtitled text involved a higher level of cognitive

effort. In the vast majority of instances, LLPs dwelled on the subtitles two or even three times more than HLPs especially when ironic meaning was modified in the subtitled text. The fact that LLPs relied on the subtitles more heavily is linked to their lower proficiency in English, as they were not able to retrieve the intended meaning from the original dialogue. As a result, subtitles were more informative for them. This evidence contradicts previous studies by d'Ydewalle et al. (1987), who found that the viewers' familiarity with the language spoken in the English video did not have an impact on the time that they spent reading the subtitles. Although LLPs spent more time reading the subtitles in the majority of instances of multimodal irony, they also examined other parts of the screen in a more exploratory way in search of visually salient elements for irony comprehension. Longer fixations on certain subtitles can also be the result of the layout of subtitles in specific parts of the scene that is not linked to the increased cognitive effort. On the whole, on the one hand, LLPs dwelled at the subtitles much longer than other participants, as anticipated, but on the other hand, they also spent much time exploring other parts of the screen in the search for additional pieces of narrative information. Although their fixations durations were not found to be statistically significant, these results enhance our understanding as to how the participants with low level of linguistic abilities in English engage with translated film tests. Eye-tracking movements did not vary significantly across the three groups in the voiced-over version, as predicted. Counter-intuitively, however, all responders remained focused predominantly on the AOIs around the characters' faces. These gaze patterns strengthen the idea that the on-screen characters' relay visually salient information to the narrative. The presence of the voice-over narrator is another plausible explanation. Since the monotonous *lektor*'s voice drowns most of the original soundtrack, several participants may have had a hard time to grasp intonation patterns, volume or pitch of voice which have been found to play an auxiliary role in the process of interpretation of attitudes of disapproval. While the qualities of characters' voices were not easily accessible due to the drowning effect of the voice-over narration, the participants looked longer at the faces in an attempt to recognise the on-screen characters' emotions and thoughts conveyed visually. To my surprise, however, the participants in this study avoided looking at the AOIs around the on-screen characters' bodies. Other parts of the screen were also largely ignored, although the participants were not distracted by the presence of subtitles at the bottom of the screen. The present results thus corroborate Treuting's (2006) and Redmond's et al. (2015) findings, i.e., that the protagonists' faces are the strongest points of attraction of visual attention, which is highly significant as facial expressions have come across as helpful in the reception

of dissociative attitude. The evidence from the eye-tracking data analysis thus confirmed my assumptions that all viewers exhibited similar visual patterns, irrespective of their level of English proficiency, in the voiced-over clips featuring the use of multimodal irony.

Given the current exposure to English and its globalized presence in the multilingual Europe, it can be assumed that the source language content is not completely foreign (Pym, 2013). Although the viewers may not be fully familiar with English, they may be relatively acquainted with it due to the fact that the overwhelming majority of films comes from English-speaking countries. Nevertheless, even if the audiences managed to read the subtitles, it does not mean they understood the intended meaning. The challenges start when more culture-specific concept like irony come into play that requires not only some linguistic knowledge of SL but also a certain level of source culture awareness to be able to spot non-verbal resources like a tone of voice or a facial expression that can facilitate the comprehension process.

4. To what extent are Polish viewers, representing different levels of English proficiency, able to grasp the intended ironic meaning in the subtitled and voiced-over Sherlock Holmes films, as shown by questionnaire responses?

The questionnaire component of my analysis complemented the analysis of eye-tracking information and aimed to explain why (i) participants sometimes looked at certain AOIs longer than expected, (ii) whether longer fixations are related to the cognitive load involved in the processing of the intended ironic meaning, and (iii) whether there is a correlation between the level of proficiency in English and the irony comprehension scores. On the whole, the qualitative and quantitative analysis of questionnaire responses brought up major differences in the retrieval of the intended meaning across many viewers with high, medium and low level of English in the subtitled and voiced-over clips. In some cases, the findings turned out to clash with my predictions. Some instances of multimodal irony which it had been assumed would be easily recoverable posed interpretative challenges, and many participants managed to access the intended ironic meaning in cases which I had considered impenetrable.

On average, the comparison of the mean comprehension scores revealed that in the overwhelming majority of the instances, irony comprehension was significantly smoother for HLPs and MLPs than for LLPs. As the eye-tracking data indicated, HLPs and MLPs largely skipped a great number of subtitles and still understood the instances of multimodal

irony to a large extent, as their responses indicate. Moreover, in the cases of the modified subtitles that HLPs and MLPs including LLPs fixated on longer than expected (SH2_1), they did not encounter major problems in identifying manifestations of the attitude of dissociation. As far as the subtitles with the verbatim content are concerned (SH1_3 and SH2_3), HLPs' and MLPs' longer fixations on the subtitled area did not have an impact on the comprehension of the intended meaning either – as the relatively high comprehension scores indicate. It's more likely that these participants spent time looking at the subtitles out of curiosity, e.g., to compare the source text with its target counterpart; they could have also looked at them automatically because they were located close to the on-screen character's face and body within medium close-up and close-up shots – thus attracting more visual attention. In addition, in those cases where the subtitler opted for either modification or preservation of the original dialogue in the subtitles, HLPs and MLPs achieved relatively high degree of comprehension. Clip SH1_1 with modified subtitles, on the other hand, has proven challenging to some LLPs to understand as well as some including these instances of multimodal irony where the original dialogue preserved in TL. Longer fixation time in cases where the ironic meaning was faithfully rendered (SH1_2 and SH2_2) in the subtitled are thus linked to LLPs' greater processing effort, as lower comprehension points earned by LLPs indicate. Thus, in the present study, it cannot be confirmed that the modified subtitles result in higher comprehension than verbatim subtitles, as has been found to be the case in the research of Schilperoord et al. (2005) or Moran (2012). In addition, as opposed to Smith's (2015) claim that "there is no evidence that subtitling leads to poorer comprehension", in the context of the current experiment, these results show evidence that subtitles have the impact on poorer comprehension of ironic content in the case of the participants with a lower level of English-language competence.

In the voiced-over version, LLPs understood the instances of multimodal irony more successfully than in its subtitled counterpart, as expected, although the vast majority of them experiences difficulties to the characteristic film score or the special effects like "antiphonal laughter". Only to a small extent, LLPs grasped some features of the on-screen characters' voice quality like intonation or tone of voice. Instead, they relied heavily on the information conveyed visually and recognised a number of non-verbal semiotic resources that belong to mise-en-scène. As a result, the participants with low level of proficiency in English retrieved multimodal irony largely via the verbal, visual and kinesic modes that were for them easier to access. Surprisingly, however, HLPs and MLPs did not spot as many non-verbal acoustic cues, as hoped. Although some of them mentioned some peculiarities of voice quality as

well as laughter, more importantly "antiphonal laughter", in the questionnaire responses, almost all of them missed the music score, in contrast to the subtitled excerpts. Hence, the voice-over narration reduced the audibility of the original soundtrack, and thereby hindered the perception of the non-verbal acoustic information. Particularly, in these clips where the film score or the on-screen character's intonation have been found to be vital in the detection of the attitude of dissociation, HLPs and MLPs provided less accurate responses in the voice-over than in the subtitled version.

On the whole, there is a strong positive correlation between comprehension scores and language groups for the SUB_VO group and the VO_SUB group, as anticipated. That is to say, the level of proficiency in English does play a significant role in the comprehension of multimodal irony, as the comprehension scores increase along with the language level.

5. What type of film translation, subtitling or voice-over, proves to be the most optimal in the reception of multimodal irony by Polish audience representing different levels of English proficiency?

As shown on several occasions in the present study, the overwhelming majority of the viewers, particularly HLPs and MLPs provided a substantial evidence base that they understood meaning, and more importantly ironic meaning, relayed with subtitles and voice-over vary to a large extent. Sometimes their irony comprehension scores were even better in the subtitled clips than in its voiced-over counterparts. In contrast to the voiced-over versions of the two Sherlock Holmes films, several participants also noticed a wide array of non-verbal acoustic semiotic resources in addition to the kinesic and visual modes in the subtitled version. HLPs and MLPs have also been found to split their visual attention between reading the subtitles and following the action in the middle of the screen rather effortlessly. Thus, subtitling is more optimal for the viewers with a high and medium level of English proficiency. When it comes to the participants whose English-language command is rather low, voice-over is a more suitable form of audiovisual translation. As the findings from the current project indicate, on average, LLPs retrieved ironic meaning better with voice-over than subtitles. It was indeed more challenging for LLPs to process the subtitled and the visual content at the same time due to their low proficiency in the source language. On a positive note, however, LLPs also demonstrated their understanding of multimodal irony in 3 out of 6 subtitled clips, which provides a good prognosis about a future use of subtitles by this group of Polish viewers.

Now let us return to the central research question in the present research project. The afore-discussed findings overall extend our knowledge of the contribution of the verbal and non-verbal modes to the reception of multimodal irony and clearly indicate that non-verbal modes are instrumental in the comprehension process of dissociative attitude in the two Sherlock Holmes films. It has also been shown in the current study that the level of immersion in the foreign language and culture and to some extent the modality of audiovisual translation determined the extent to which viewers grasped the non-verbal semiotic resources and interpreted their meanings associated with certain types of emotions and attitudes. Although some similarities have also been observed between the SUB_VO and VO_SUB groups and within the language subgroups, the aforementioned results corroborate the common opinion in the film studies that there is no identical response to the film text as audiences follow their own preferred way of watching films in respect to their abilities and needs (Hall et al., 2013; Wharton and Grant, 2005; Stafford 2007).

Some complementary conclusions have also been formulated from the experimental study in reference to the perception of the non-verbal semiotic resources in the comprehension process that intended to enhance our understanding of irony recovery from a Cultural Studies perspective. It has been mentioned on several occasions throughout this thesis that irony is a highly culture-specific phenomenon. The results of this study suggest that there are some non-verbal visual cues linked to features of facial expressions such as "a blank face", raised eyebrows, wide opened eyes or smirking that convey similar meanings to Polish and American audiences in the construal of multimodal irony. As a result, sometimes ironic meaning is construed by the non-verbal cues that may be understood in the source culture and beyond the boundaries of that culture. More empirical studies would be necessary, however, to provide some solid evidence for comprehension of multimodal irony in the cross-cultural context.

8.2 Limitations of the Research

Although it has been endeavoured to be as meticulous as possible in the design the current study, this study has a number of limitations that may affect the generalisability of the findings.

The first set of limitations concerns the type of audiovisual material selected for the current study. The use the two Sherlock Holmes films directed by Guy Ritchie (2009 and 2011) might have provided my participants in the experimental part of my analysis with some clues as to the type of phenomenon that it has been chosen to investigate. Although the participants were only given a vague description of the nature and purpose of my

experiment, the presence of Robert Downey Jr. starring as Sherlock Holmes might have given some hints suggesting that the actual goal of the study was the perception of irony or humour. Using a selection of clips from different film texts or genres would have helped to manage the viewers expectations, in the sense of concealing what they would be asked for in the experiment – although selecting different excerpts might also pose additional interpretative challenges. To address this limitation of my study, I endeavoured to include instances of multimodal irony from the Sherlock Holmes films in which the attitude of dissociation was shown not solely by Sherlock, but also by other protagonists in the film like Watson or Mycroft.

Another set of limitations refers to the experimental study. In this part of the study it has been decided to draw on the eye-tracking technology, as it provides a very precise logging of the gaze position on the screen. In the analysis of the eye-tracking data, the three main measurements of eye movement have been selected, that is, the number of fixations, the total fixation duration, and the mean fixation duration, which provided me with reliable information to establish where the target audience placed their visual attention on the screen and assess how much cognitive effort they needed to invest in order to process the intended meaning. There are, however, several other eye-tracking metrics applied in the other AVT studies (Kruger and Steyn, 2014) that could have been beneficial to the current study. Other eye-tracking measures were not incorporated in the present study because of two main reasons. First, in the current technological state eye-tracking software is more suitable for the analysis of static texts (e.g., pictures) than dynamic texts like films, for instance, to extract more detailed information about reading behaviour of the subtitled texts. In this case, the processing and analysis of more eye-tracking metrics should have been done manually, which would have been both a challenging and time-consuming task. Second, the speed of the eye tracker also determines the types of metrics that be employed. The current experiment was recorded with a sampling rate of 120Hz which allows to use fixation-based metrics measuring the position of the eye. In order to incorporate saccadic-based metrics measuring the movement of the eye, a sampling rate of 250Hz (minimum) is required. Since the current study was recorded with a sampling rate of 120Hz the application of saccadic-based measurement was automatically excluded from the present study.

8.3 Originality of this Study

Despite the aforementioned limitations, the present research project effectively enhances our understanding of the construal, relay and reception of multimodal irony in the two Sherlock Holmes films from an interdisciplinary approach.

First, the current study contributes to the mainstream of cognitive pragmatics research that so far has concentrated on the production and reception of irony in written or pre-recorded spoken dialogues. This thesis elaborates on the construal of irony in the complex multimodal environment of voiced-over and subtitled films in which the verbal and non-verbal modes co-deploy in the production of ironic meaning on screen. Additionally, the experimental and interactionist descriptive component of this research project gathered empirical evidence on the recovery of multimodal irony by the target audiences with high, medium and low levels of proficiency in the source language. It is particularly valuable in the light of the Polish audiovisual scene to learn how the Polish consume on-screen information with voice-over and subtitles and when they experience difficulties to retrieve the intended meaning.

Second, apart from cognitive pragmatics, the current research project also hopes to have contributed to the area of film studies through my analysis of the verbal, visual, kinesic and acoustic modes of the language of film and the way in which they are used to construe multimodal irony. Drawing on the adapted model of multimodal concordance, the non-verbal resources pertaining to mise-en-scène, cinematography, editing and soundtrack that composed the ironic meaning have been singled out. With the use of eye-tracking technology and questionnaires, the information from participants was extracted to see whether they recognised the filmmaker's intentions and whether the non-verbal resources attracted their attention. Since "irony comprehension is a complex cognitive task" (Hala et al., 2010: 299) the filmmakers should not take for granted that viewers are (or not) always able to identify and understand their visions on screen.

Third, the current thesis also delivers an obvious contribution to the area of AVT: It represents the first step towards the examination of the multimodal nature of irony not only in the subtitled but also in the voiced-over film text. The results obtained from this empirical study of irony recovery can hopefully support and facilitate audiovisual translators' decision-making processes in the future. Being aware of how viewers with a high, medium and low level of English retrieve ironic meaning with subtitles and voice-over also constitutes a valuable source of information to make audiovisual translators more sensitive about the

complexity of film semiotics and pragmatics and stimulate them to consider how they can accommodate the needs of the vast majority of audiences (cf. Karamitroglou, 2006).

Finally, it is hoped that this thesis represents a step towards a considerable change in the future redesign of the audiovisual landscape on television in Poland. As the empirical part of this study clearly indicates, viwers with high and medium level of English were able to retrieve the complex phenomenon of multimodal irony with subtitles and voice-over rather effortlessly. Even the some subjects possessing low level of English proficiency were quite successful in understanding multimodal irony in half of the instances featured in my chosen subtitled and voiced-over clips. What is more, despite the fact that the voice-over narration still prevails on the digital terrestrial television in Poland, the recent shift towards subtitles is clearly visible within the audiovisual market. Given the fact that Polish viewers in this project overall process even complex linguistic phenomenon like irony in the film text even more successfully with subtitles than with voice-over, members of the audiovisual industry should be more inclined to introduce this modality as an optional form of audiovisual translation to meet the needs of these viewers. The availability of subtitles on TV in Poland would not only benefit the participants with a good command of the second language but also it is also very likely that their presence on a daily basis on TV would encourage the viewers with a low level of English to enhance their linguistic skills. Furthermore, subtitling is currently favoured by the majority of Polish audiences (Szarkowska and Laskowska, 2014). This strong trend towards subtitling is also confirmed by the participants' responses in the current study. That is, 76% of the ones surveyed expressed their preference towards subtitles and only 24% towards voice-over watching the subtitled films a few times a month and voiced-over films a few times a year. On the top of this, digital technology is developing now at remarkable speed and people will be exposed to even more complex multimodal and dynamic texts in the near future, for instance in the form of augmented or virtual reality. It would be therefore highly advantageous for the target audiences, and more importantly for these viewers whose level of English is rather elementary, to familiarise themselves with this type of "fleeting text on screen" on a regular basis. These findings should thus provide members of the audiovisual industry in Poland with sufficient evidence to introduce subtitles as an additional translation option in the digital age.

8.4 Suggestions for Further Research

The present research project constitutes the first attempt to explore the construal, relay and comprehension of multimodal irony in the film text using a triangulation of methods. The findings of this study not only add valuable insights into the field of AVT, but also open new avenues for further research.

The design of the methodological apparatus has been adapted to cater for the analysis of multimodal irony in subtitled and voiced-over films. Similarly, it can also be successfully used for the study of dubbed films, other film genres and other language combinations. It would be particularly interesting to explore the attitude of dissociation in other film genres and language pairs. Comparing irony comprehension between source and target audiences would also generate interesting findings, especially from the perspective of cultural studies. Specifically, it would be interesting to see if source and target viewers have (or have not) common understanding of non-verbal resources (e.g., facial expressions) that construe multimodal irony on screen.

The findings of this research provide useful insights for filmmakers, audiovisual translators, scholars and members of the AVT industry pertaining to the ways in which viewers with different levels of English proficiency process and make sense of complex meaning that is construed multimodally. It would be also worth investigating the relationship between irony comprehension and other audience characteristics like age, gender or personality traits. For instance, in the context of Polish AVT scene it would be interesting to examine the age group in their 50s and 60s who are used to watching films with voiced-over to see whether and/or to what extent they would be able to retrieve the intended meaning in subtitled films.

Another potential research avenue would involve examining the impact of emotional intelligence on the recognition of emotions associated with transfer of ironic meaning. The concept emotional intelligence refers to our capabilities to express and perceive emotions. The study of emotional awareness has been brought to the fore in the foreign language education and attitudes and emotions have become an important part of foreign language acquisition. The current findings highlight the link between emotion recognition, e.g., in facial expressions and viewers' level of immersion in the foreign language and culture. Emotions are generally construed by the culture we live in and the language we communicate with (e.g., Ożańska-Ponikwia, 2013), so the interpretation of emotions may pose interpretative challenges for speakers of other languages. It would be interesting to explore the complex relationship between emotional intelligence and perception and/or expression

of emotions like irony in the first and second language in much more detail both in face-to-face interactions and in the multimodal texts like films.

References

Abrams, M.H. (1993) *A Glossary of Literary Terms,* Texas: Harcourt Brace Jovanovich College Publishers.

Ackerman, B. (1983) 'Form and Function in Children's Understanding of Ironic Utterances', *Journal of Experimental Child Psychology* (35) 487–508. Retrieved from http://www.sciencedirect.com/science/article/pii/0022096583900231#

Adachi, T. (1996) 'Sarcasm in Japanese', *Studies in Language* 20 (1): 1–36. Retrieved from https://www.researchgate.net/publication/233508544_Sarcasm_in_Japanese

Anolli, L., Ciceri, R. & Infantino, M.G. (2002) 'From 'Blame by Praise' to 'Praise by Blame': Analysis of vocal patterns in ironic communication', *International Journal of Psychology* (37): 266–276. Retrieved from http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=4197853d-9dee-473c-b4dc-48913dd7ddb4%40sessionmgr113&vid=1&hid=124

Anolli, L., Ciceri, R. & Infantino, M., G. (2000) 'Irony as a Aame of Implicitness: Acoustic profiles of ironic communication', *Journal of Psycholinguistic Research* 29 (3): 275–311. Retrieved from https://link.springer.com/article/10.1023/A:1005100221723

Adams, A. & Cox, A.L. (2008) 'Questionnaires, In-depth Interviews and Focus Groups', in: Cairns, P. & A.L. Cox, A., L. (eds) *Research methods for human computer interaction,* Cambridge: Cambridge University Press, 17–34.

Antonijevic, S. (2008) *Rhetoric, Culture and Avatars: A microethographic analysis of nonverbal communication in multiuser 3D virtual environment* (Doctoral dissertation), University of Minnesota, USA.

Attardo, S., Eisterhold, J., Hay, J. & Poggi, I. (2003) 'Multimodal Markers of Irony and Sarcasm', *International Journal of Humor Research* (16): 243–260. Retrieved from http://www.degruyter.com/view/j/humr.2003.16.issue-2/humr.2003.012/humr.2003.012.xml?format=INT

Baldry, A. & Thibault, P. (2006), *Multimodal Trancription and Text Analysis. A Multimodal Toolkit and Coursebook*, London: Equinox.

Baker, M. (1992), *In Other Words: A Coursebook on Translation*, London: Routledge.

Barbe, K. (1995) *Irony in Context*, Amsterdam: John Benjamins Publishing.

Beaver, E., F. (2007) *Dictionary of Film Terms: The Aesthetic Companion to Film Art*, New York: Peter Lang Publishing.

Boggs, M., J. and Petrie, D. W. (2008) *The Art of Watching Films*, New York: McGraw-Hill.

Bogucki, Ł. (2004) 'The Constraint of Relevance in Subtitling', *Journal of Specialised Translation* (1): 69–85. Retrieved from www.jostrans.org/issue01/art_bogucki_en.pdf

Bogucki, Ł. & Kredens, K. (2010), *Perspectives on Audiovisual Translation*, Frankfurt am Main: Peter Lang.

Bojko, A. (2009) 'Informative or Misleading? Heatmaps Deconstructed', in: *Human-Computer Interaction. New Trends,* Berlin: Springer, 30–39.

Booth, W. (1984) *A Rhetoric of Irony*, Chicago: Chicago University Press.

Bordwell, D. & Thomson, K. (1997) *Film Art,* New York: McGraw-Hill.

Bryant, G. & Fox Tree, J. E. (2002) 'Recognising Verbal Irony in Spontaneous Speech', *Metaphor and Symbol* (17): 99-117. Retrieved from http://www.gregbryant.org/mandsym.pdf

Bryant, G. A. & Fox Tree, J. E. (2005) 'Is There an Ironic Tone of Voice?', *Language and Speech* 48(3); 257–277. Retrieved from https://www.researchgate.net/profile/Jean_Fox_Tree/publication/7355390_Is_there_an_Ironic_Tone_of_Voice/links/0c9605252d309de9ad000000/Is-there-an-Ironic-Tone-of-Voice.pdf

Bryant, G. A. (2011) 'Verbal Irony in the Wild', *Pragmatics and Cognition* 19 (2): 291–309. Retrieved from www.gregbryant.org/ironyinthewildPC2011.pdf

Bucaria, C. & Chiaro, D. (2007) 'End-user Perception of Screen Translation: The case of Italian dubbing', *TradTerm* (13): 91–118. Retrieved from http://myrtus.uspnet.usp.br/tradterm/site/images/revistas/v13n1/v13n1a06.pdf

Bucaria, C. (2005) 'The Perception of Humour in Dubbing vs Subtitling: The Case of *Six Feet',* ESP Across Cultures (2): 34–46. Retrieved from http://edipuglia.it/wp-content/uploads/2015/05/ESP_2_2005.pdf

Burr, T. (2011), 'Downey is right at Holmes'. Retrieved from http://archive.boston.com/ae/movies/articles/2009/12/25/downey_is_right_at_holmes/

Caffrey, C. (2009) *Relevant Abuse? Investigating the Effects of an Abusive Subtitling Procedure on the Perception of TV Anime Using Eye Tracker and Questionnaire* (Doctoral dissertation). Dublin City University, Dublin, Ireland.

Chakhkachiro, R. (2009) 'Translating Irony in Political Commentary Texts From English into Arabic'*, Meta: Translators' Journal* 54 (1): 216–240. Retrieved from http://id.erudit.org/iderudit/029792ar

Chandler, D. (2002) *Semiotics: The Basics*, London: Routledge.

Chandler, G. (2009) *Film Editing: Great Cuts Every Filmmaker and Movie Lover Must Know*, Studio City: Michael Wiese Productions.

Chatman, S., B. (1978) *Story and Discourse: Narrative Structure in Fiction and Film*, New York: Cornell University Press.

Chaume, F. (1997) 'Translating Non-verbal Information in Dubbing', in: Poyatos, F. (ed.) *Non-Verbal Communication and Translation. New Perspectives and Challenges in Literature, Interpretation and the Media*, Amsterdam: John Benjamins, 315–326-

Chaume, F. (2013) 'The Turn of Audiovisual Translation. New Audiences and new Technologies', *Translation Spaces* (2): 107–125.

Clark, B. (2013) *Relevance Theory*, Cambridge: Cambridge University Press.

Clark, H., & Gerrig, R., (1984) 'On the Pretense Theory of Irony', *Journal of Experimental Psychology* (113): 121–126. Retrieved from http://web.stanford.edu/~clark/1980s/Clark,%20H.H.%20_%20Gerrig,%20R.J.%20_On%20the%20pretense%20theory%20of%20irony_%201984.pdf

Coutrot, A., Guyader N., Ionesc, G. & Caplier., A. (2012) 'Influence of Soundtrack on Eye Movements During Video Exploration', *Journal of Eye Movement Research 5*(4.2): 1–10.

Cummings, L. (2010) *The Pragmatics Encyclopedia,* London: Routledge.

Currie, G. (2005) 'Why irony is pretence', in: Nichols, S. (ed.) *The Architecture of the Imagination,* Oxford: Oxford University Press, 122–138-

Dancyger, K. (2013) *The Technique of Film and Video Editing: History, Theory, and Practise*, Burlington: Elsevier.

Desilla, L. (2009) *Towards a Methodology for the Study of Implicatures in Subtitled Films: Multimodal Construal and Reception of Pragmatic Meaning Across Cultures* (Doctoral dissertation). University of Manchester, Manchester, UK.

De Wilde, J. (2010) 'The Analysis of Translated Literary Irony: Some Methodological Issues', *Linguistica Antwerpiensia*: *Translating Irony* (9): 25–44. Retrieved from https://lanstts.uantwerpen.be/index.php/LANS-TTS/article/view/260/162:

Díaz Cintas, J. (2001) *La Traducción Audiovisual: el Subtitulado*, Salamanca: Ediciones Almar.

Díaz Cintas, J. & Remael., A. (2007) *Audiovisual Translation: Subtitling*, Manchester: St. Jerome.

Díaz Cintas, J. & Fernandez Cruz, M. (2008) 'Using Subtitled Video Materials for Foreign Language Instruction', in: J. Díaz Cintas (ed.) *The Didactics of Audiovisual Translation,* Amsterdam: John Benjamins, 201–214.

Di Giovani, E., (2003) 'Cultural Otherness and Global Communication in Walt Disney Films at the Turn of the Century', *The Translator* (9): 207–223. Retrieved from www.tandfonline.com/doi/abs/10.1080/13556509.2003.10799154

Dix, A. (2008) *Beginning Film Studies*, Manchester: Manchester University Press.

Duchowski, A. (2007) *Eye tracking methodology: Theory and practice,* New York: Springer Science & Business Media.

Dries, J. (1995) 'Breaking Eastern European Barriers', *Sequentia* 2 (4), 6–7.

Dyer, A., G. & Pink, S. (2015) 'Movement, Attention and Movies: The Possibilities and Limitations of Eye Tracking?', *Refractory: Journal of Entertainment Media* 25. Retrieved from http://refractory.unimelb.edu.au/2015/02/07/dyerpink

Dyer, R. (1998) *Stars,* London: BFI.

Dyer, R. (2007) *Pastiche,* New York: Routledge.

Dynel, M. (2014), 'Linguistic approaches to (non) humorous irony', *Humour* 27 (4): 537–550.

Dwyer T. (2015) 'From Subtitles to SMS: Eye Tracking, Texting and Sherlock', *Refractory: Journal of Entertainment Media* 25. Retrieved fromhttp://refractory.unimelb.edu.au/2015/02/07/dwyer/

d'Ydewalle, G., Van Rensbergen, J., & Pollet, J. (1987) 'Reading a Message when the Same Message Is Available Auditorily in Another Language: The case of subtitling', in: J.K O'Reagan & A. Lévy Schoen (eds) *Eye Movements: From Physiology to Cognition,* Amsterdam: Elsevier Science Publishers B.V, 313–321.

d'Ydewalle, G. & De Bruycker, W. (2007) 'Eye Movements of Children and Adults while Reading Television Subtitles', *European Psychologist* (12): 196–205. Retrieved from http://academic.csuohio.edu/kneuendorf/frames/subtitling/dYdewalle&DeBruycker.2007.pdf

Elleström, L. (2002) *On interpreting Literature, Music and the Visual Arts Ironically*, New York: Rosemont Publishing & Printing Corp.

Evans, D. (2001) *Emotion: A Very Short Introduction,* London: Oxford University Press.

Feltrin-Morris, M. (2011) 'The Stuff Irony Is Made of: Translators as scholars', *Linguistica Antwerpiensia*: *Translating Irony* (9): 213-231. Retrieved from https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/269/171

Fónagy, I. & Magdics, K. (1963) 'Emotional Patterns in Intonation and Music', *Zeitschrift fur Phonetik* (16): 293-326.,.

Franco, E., Matamala, A. & Orero, P. (2010) *Voice-over Translation: An Overview*, Bern: Peter Lang.

Fuentes Luque, A. (2003) 'An Empirical Approach to the Reception of AV Translated Humour', *The Translator* (9): 293–306. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/13556509.2003.10799158

Gallaghter, M. (2006) *Action Figures: Men, Action Films, and Contemporary Adventure Narratives,* New York: Palgrave Macmilian.

Gambier, I. (2003) 'Screen Transadaptation: Perception and Reception', *The Translator* (9): 171–189. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/13556509.2003.10799152

Garrand, T., P. (1984) *The Comedy Screenwriting of Preston Sturges: An Analysis of Seven Paramount Auteurist Screenplays* (Doctoral dissertation). University of Southern California, USA.

Garcarz, M. (2007) *Przeklad slangu w filmie. Telewizyjne przeklady filmow amwrykanskich na jezyk polski.* [Translating slang in films. Television translations of American films into Polish], Krakow.

Georgakopoulou, P. (2010) *Reduction Levels in Subtitling. DVD Subtitling: A Convergence of Trends,* Saarbrücken: Lambert Academic Publishing.

Ghia, Elisa. (2012) 'The Impact of Translation Strategies on Subtitle Reading', in: E. Perego (ed.) *Eye Tracking in Audiovisual Translation,* Rome: Aracne Editrice, 155–182.

Gibbs, R.W & O'Brien, J. (1991) 'Psychological Aspects of Irony Understanding', *Journal of Pragmatics,* 16 (6): 523–530. Retrieved from http://ac.els-cdn.com/0378216691901013/1-s2.0-0378216691901013-main.pdf?_tid=0c94b860-95f7-11e4-a390-00000aacb362&acdnat=1420585082_b10a8a0548e988d0b3ec88704fcb728a

Gibbs, R. W. (1994) *The Poetics of Mind, Figurative Thought, Language and Understanding*, New York: Cambridge University Press.

Gibbs, R. & Colston, H. (2007) *Irony in Language and Thought*, New York: Taylor & Francis Group.

Gibbs, R. W. (2011) 'Irony in Talk Among Friends', *Metaphor and Symbol* 15 (1-2): 5–27. Retrieved from https://dx.doi.org/10.1080/10926488.2000.9678862

Gottlieb, H. (2004) 'Language-political Implications of Subtitling', in: P. Orero (ed.) *Topics in Audiovisual Translation,* Philadelphia: John Benjamins Publishig Co, 83–100.

Gottlieb, H. (2005) 'Multidimensional Translation: Semantics turned Semiotics', *MuTra 2005 – Challenges of Multidimensional Translation: Conference Proceedings.* Retrieved from http://www.euroconferences.info/proceedings/2005_Proceedings/2005_Gottlieb_Henrik.pdf

Grice, H.P. (1967/75) 'Logic and Conversation', *Syntax and Semantics: Speech acts* 3): 41–58. Retrieved from http://www.ucl.ac.uk/ls/studypacks/Grice-Logic.pdf

Grice, H.P. (1978/89) 'Further Noted on Logic and Conversation', in: *Studies in the Way of Words,* Cambridge: Harvard University Press, 41–58. Retrieved from: http://aardvark.ucsd.edu/language/grice_further_notes.pdf

Grundy, P. (2000) *Doing Pragmatics,* London: Rutledge.

Gurilli, L., R. & Ortega, A., B., M. (2013) *Irony and Humour. From Pragmatics to Discourse,* Amsterdam: John Benjamin.

Gutt, E., A. (1991) *Translation and Relevance. Cognition and Context,* Oxford: Blackwell.

Haiman, J. (1990) 'Sarcasm as Theater', *Cognitive Linguistics* 1 (2): 181–201. Retrieved from https://www.degruyter.com/view/j/cogl.1990.1.issue-2/cogl.1990.1.2.181/cogl.1990.1.2.181.xml?rskey=IUZcDV

Haiman, J. (1997) *Talk is cheap: Sarcasm, Alienation and the Evolution of Language,* Oxford: Oxford University Press.

Hala, S., Pexman, P., Cimie E. Rostd, K. & Glenwright M. (2010) 'A Bidirectional View of Executive Functions and Social Interaction', in: S. Bryan, U. Mueller, J. Carpendale, A. Young, G. Iarocci (ds) *Self and Social Regulations. Social Interaction and the Development of Social Understanding.,* New York: Oxford University Press.

Hall, S. (1997) *Representation: Cultural Representations and Signifying Practices,* London: Sage.

Hall, H. R., Lockwood, N. & Sheng, H., (2013) 'Psychological Assessment Tools for the Study of Evaluation of Learning Technologies', in: S. Yamamoto (ed.) *Human Interface and the Management of Information*, New York: Springer.

Hamamoto, H. (1998) 'Irony from a Cognitive Perspective', in: R. Carston & S. Uchida (eds) *Relevance Theory: Applications and Implications,* Amsterdam: John Benjamins Publishing, 257–270.

Hatim, B. & Masón, I. (1990) *Discourse and the Translator,* London: Longman.

Hatim, B. (1997) *Communication across Cultures. Translation Theory and Contrastive Text Linguistics,* Exeter: University of Exeter Press.

Hayword, S. (2013) *Cinema Studies: The Key Concepts,* New York: Routledge.

Herdández, N., Aude, M., Rémy, M., Bonnet-Brilhault, M., Roux, F., Barthelemy, S. & Joëlle S. (2009) 'Exploration of Core Features of a Human Face by Healthy and Autistic Adults Analyzed by Visual Scanning', *Neuropsychologia* 47 (4):1004–1012. Retrieved from https://www.deepdyve.com/lp/elsevier/exploration-of-core-features-of-a-human-face-by-healthy-and-autistic-OkRlB0f8W9

Hershler, O. & Hochstein, S. (2005) 'At first sight: a high-level pop out effect for faces', *Vision Research* (45): 1707–1724. Retrieved from https://www.sciencedirect.com/science/article/pii/S0042698905000337/pdfft?md5=514d559510abbf4245f5441918bf736c&pid=1-s2.0-S0042698905000337-main.pdf

Hoesterey, I. (2001) *Pastiche Cultural Memory in Art, Film, Literature,* Indiana: Indiana University Press.

Holdcroft, D. (1983) 'Irony as a Trope, and Irony as Discourse', *Poetics Today* 4 (3): 493–511.

Holmqvist, K., Nystrom, M. & Andersson, R. (2011) *Eyetracking. A Comprehensive Guide to Methods and Measures,* Oxford: Oxford University Press.

Hołobut, A. (2015) 'Voiceover as Spoken Discourse', in: J. Díaz Cintas & R. Baños Piñero (eds) *Audiovisual Translation in a Global Context. Mapping an Ever-changing Landscape,* London: Palgrave Macmillan, 225–252.

Holsanova, J. (2014a) 'Reception of Multimodality: Applying eye tracking methodology in multimodal research', in: C. Jewitt (ed.) *Routledge Handbook of Multimodal Analysis,* London: Routledge, 285–296.

Hutcheon, L. (1994) *Irony's Edge*: *The Theory and Politics of Irony,* London: Routledge.

Hvelplund, K., T. (2014) 'Eye Tracking and the Translation Process: Reflections on the analysis and interpretation of eye tracking data', .in: R. Muñoz (ed.) *Minding Translation* (1): 201–223. Retrieved from http://www.e-revistes.uji.es/index.php/monti/article/view/1706

Jacobson, R. (1959/2004) 'On linguistics aspects of translation', in; L., Venuti (ed.) *The Translation Studies Reader,* New York: Routledge, 138–143.

Jefferson, G. (1984) 'On the Organization of Laughter in Talk about Troubles', in: J., M. Atkinson & J. Heritage (eds) *Structures of Social Action,* Cambridge: Cambridge University, 346–369.

Jensen, L. (2014) Representations of Sherlock Holmes. A Study of Sherlock Holmes' New Trademarks as a Sex symbol, Action hero, and Comedian in the 21st Century. Master thesis submitted to Aalborg University.

Jensema, C. (2000) 'Eye Movement Patterns of Captioned TV Viewers', *American Annals of the Deaf* 145 (3): 275–285. Retrieved from https://www.researchgate.net/publication/12355679_Eye_Movement_Patterns_of_Captioned_Television_Viewers

Karamitroglou, F. (2006) 'Look, Listen, Learn: Ideal courses in audiovisual translation training'. Paper presented at the *Translation in the University and the Workplace International Symposium*, organised by the Department of Translation & Intercultural Studies of the Aristotle University of Thessaloniki, Greece, 3–4 November 2006.: Retrieved from www.enl.auth.gr/events/karamitroglou.ppt

Karoń, B. (2006) 'Zmiany w systemie edukacji po 1989 roku'. Retrieved from http://www.lewica.pl/?id=11075

Kaufer, D., S. (1981) 'Understanding Ironic Communication', *Journal of Pragmatics,* 5 (6): 495–510. Retrieved from https://www.sciencedirect.com/science/article/pii/0378216681900151

Kellison, C. (2012) *Producing for TV and Video: A Real-World Approach,* New York: Elsevier.

Klinger, B. (2006) *Beyond the Multiplex: Cinema, New Technologies and the Home*, Berkeley: University of California Press.

Kotelecka, J. K. (2006) 'Traducción Audiovisual para la Televisión en Polonia: Versión locutada. Entrevista con Barbara Rodkiewicz-Gronowska (TVP) [Audiovisual translation for television in Poland: voice- over. An interview with Barbara Rodkiewicz-Gronowska (TVP)]. *Trans* (10): 157–168.

Kozloff, S. (2000) *Overhearing Film Dialogue*, California: University of California Press.

Kress, G. & van Leeuwen, T. (1996) *Reading Images. The Grammar of Visual Design*, London: Routledge.

Kreuz, R. J., & Roberts, R. (1995) 'Two Cues for Verbal Irony: Hyperbole and the ironic tone of voice', *Metaphor and Symbolic Activity* 10 (1): 21–31. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/s15327868ms1001_3

Kruger, J–L. (2012) 'Making Meaning in AVT: Eye tracking and viewer construction of narrative', *Perspectives: Studies in Translatology* 20 (1): 67–86. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/0907676X.2011.632688

Kruger, J–L., Mattew, G. & Hefer, E. (2013) 'Measuring the Impact of Subtitles on Cognitive Load: Eye tracking and dynamic audiovisual texts', *Proceedings of the 2013 Conference on Eye Tracking South Africa* (1): 62–66. Retrieved from https://dl.acm.org/citation.cfm?id=2509331

Kruger, J–L. & Steyn, F. (2014) 'Subtitles and Eye Tracking: Reading and Performance', *Reading Research Quarterly* 49 (1): 105–120. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/rrq.59/epdf

Kruger, J–L., Szarkowska, A., & Krejtz, I. (2015) 'Subtitles on the Moving Image: An Overview of Eye Tracking Studies', *Journal of Entertainment Media*: *Eye-Tracking the Moving Image* 25. Retrieved from refractory.unimelb.edu.au/2015/02/07/kruger-szarkowska-krejtz/

Kuhn, A. & Westwell, G. (2012) *A Dictionary of Film Studies,* Oxford: Oxford University Press.

Künzli, A. & Ehrensberger-Dow, M. (2011) ‚Innovative Subtitling: A Reception Study', in: C. Alvstad, A. Hild & E. Tiselius (eds) *Methods and Strategies of Process Research: Integrative approaches in Translation Studies,* Amsterdam: John Benjamins Publishing Company, 187–200.

Lee, M., Roskos, B. & Ewoldsen, D., R. (2013) 'The Impact of Subtitles on Comprehension of Narrative Film', *Media Psychology* 16 (4): 412–440. Rerieved from http://www.tandfonline.com/doi/pdf/10.1080/15213269.2013.826119?needAccess=true

Lyall, S. (2009) 'Is That You, Sherlock?'. Retrieved from http://www.nytimes.com/2009/01/25/movies/25lyal.html?pagewanted=all&_r=0

Macdowell, J. (2016) *Irony in Film,* London: Palgrave Macmillan.

Machin, D. (2014) 'Multimodality and Theories of the Visual', in: J., Carey (ed.) *The Routledge Handbook of Multimodal Analysis,* London: Routledge, 217–227.

Mateo, M. (1995) 'The Translation of Irony', *Meta: Translators' Journal* 40 (1): 171–178. Retrieved from http://www.erudit.org/revue/meta/1995/v40/n1/003595ar.html

Mernit, B. (2001) *Writing the Romantic Comedy*, New York: Harper Collins.

Milosky, M., L. & Ford, A. J. (1997) 'The Role of Prosody in Children's Inferences of Ironic Intent', *Discourse Processes 2*3 (1): 47–61. Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/01638539709544981

Mercado, G. (2010) *The Filmmaker's Eye: Learning (and Breaking) the Rules of Cinematic Composition,* Burlington: Focal Press.

Misiak, A. (2006) 'Politically Involved Filmmaker: Aleksander Ford and film censorship in Poland after 1945', *Cinema A Journal for Film and Audiovisual Media.* Retrieved from http://www.kinema.uwaterloo.ca/article.php?id=114&

Mirecka, I. (2002) 'Szeptanie zza ekranu'. Retrieved from https://www.tygodnikprzeglad.pl/szeptanie-zza-ekranu/

Moran, S. (2012) 'The Effect of Linguistic Variation on Subtitle Reception', in: E. Perego (ed.) *Eye Tracking in Audiovisual Translation,* Rome: Aracne Editrice, 183–222.

Murray, R. (2009) 'Sherlock Holmes'. Retrieved from http://movies.about.com/od/sherlockholmes/fr/sherlock-holmes-review.htm

Muecke, D.C. (1969) *The Compass of Irony,* London: Methuen.

Muecke, D. C. (1978) *Irony Markers*, London: Methuen.

Muecke, D. C. (1982) *Irony and the Ironist*, London: Methuen.

Nelmes, J. (2012) *Introduction to Film Studies,* New York: Routledge.

Nida, E. (1964) *Toward a Science of Translating,* Leiden: E.K. Brill.

Nilsson, J. (2013) *American Film Satire in the 1990s Hollywood Subversion*, New York: Palgrave Macmillan.

Norris, S. (2007) *Analysing Multimodal Interaction: A Methodological Framework,* London: Routledge.

O'Brien, S. (2010) 'Eye Tracking in Translation Process Research: Methodological challenges and solutions', *Copenhagen Studies in Language* (38): 251–266. Retrieved from http://doras.dcu.ie/17157/1/CSL_38_Methodologies.pdf

Orero, P. (2006) 'Synchronization in Voice-over', in: J. M. Bravo (ed.) *A new spectrum of translation studies,* Valladolid: Publicaciones de la Universidad de Valladolid, 255–264.

Orpen, V. (2003) *Film Editing: The Art of the Expressive,* New York: Wallflower.

Orrego-Carmona, D. (2014b) 'Where is the Audience? Testing the Audience Reception of Non-professional Subtitling', in: E. Torres-Simón and D. Orrego-Carmona (eds) *Tarragona: Intercultural Studies Group,* Tarragona: University of Rovira i Virgili, 77–92.

Orrego-Carmona, D. (2015) *The Reception of (Non)professional Subtitles* (Doctoral dissertation). University of Rovira i Virgii, Tarragona, Spain.

Ożańska-Ponikwia, K. (2013) *Emotions from a Bilingual Point of View. Personality and Emotional Intelligence in Relation to Perception and Expression of Emotions in the L1 and L2*, Newcastle: Cambridge Scholars Publishing.

Papenmeier, F. & Huff, M. (2010) 'DynAOI: A tool for matching eye movement data with dynamic areas of interest in animations and movies', *Behavior Research Methods,* (42): 179–187. Retrieved from
https://link.springer.com/content/pdf/10.3758%2FBRM.42.1.179.pdf

Pavlenko A., (2008) 'Emotion and Emotion-laden Words in the Bilingual Lexicon', *Bilingualism: Language and Cognition,* (11): 147–164. Retrieved from
http://www.anetapavlenko.com/pdf/BLC_2008.pdf

Pelsmaeker, K. & Van Besien, F. (2002) 'Subtitling Irony: Blackadder in Dutch', *The Translator: Translating Humour* 8 (2): 241–266. Retrieved from
http://www.tandfonline.com/doi/pdf/10.1080/13556509.2002.10799134

Perego, E., Del Missier, F., Porta, M. & Mosconi, M. (2010) 'The Cognitive Effectiveness of Subtitle Processing', *Media Psychology* 13 (3): 243–272. Retrieved from
http://academic.csuohio.edu/kneuendorf/frames/subtitling/Perego.etal.2010.pdf

Pérez-González, L. (2014) *Audiovisual Translation: Theories, methods and issues,* Abingdon: Routledge.

Pexman, P. M. & Zvaigzne, M. T. (2004)' Does Irony Go Better With Friends?', *Metaphor and Symbol*e 19 (2): 234–256. Retrieved from
http://www.tandfonline.com/doi/abs/10.1207/s15327868ms1902_3

Phillips, P. (2000) *Understanding Film Text: Meaning and Experience*, London: British Film Institute.

Pietrzyk, M. (2010) 'Zmartwychwstanie Sherlocka Holmesa'. Retrieved from
http://www.filmweb.pl/reviews/Zmartwychwstanie+Sherlocka+Holmesa-9073

Pugliese, R. (2010) 'Ironie als interkultureller Stolperstall. Grass *Beim Häuter der Zwiebel* im Spiegel der italienicher Übersetzung', *Linguistica Antwerpiensa: Translating Irony* (9): 45–63. Retrieved from https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/261

Noh, E. J. (2000) *Metarepresentation. A Relevance-Theory Approach,* Amsterdam: John Benjamins Publishing.

Ramos Pinto, S. (2013) 'How Accessible Are Audiovisual Products: A reception study of subtitled film'. Paper presented to the 5th International Conference Media for All: Audiovisual Translation, Expanding Borders. Dubrovnik.

Robinson, J., Stadler, J. & Rassell, A. (2015) 'Sound and Sight: An Exploratory Look at Saving Private Ryan through the Eye-tracking Lens', *Refractory: Journal of Entertainment Media* 25. Retrieved from
http://refractory.unimelb.edu.au/2015/02/07/robinsonstadlerrassell/

Rockwell, P. (2000) 'Lower, Slower, Louder: Vocal Cues of Sarcasm', *Journal of Psycholinguistics Research* (29): 483–495. Retrieved from
http://download.springer.com/static/pdf/271/art%253A10.1023%252FA%253A100512010
9296.pdf?auth66=1409691301_245b3aa1f931a126f0cfc6e8e910e4f9&ext=.pdf

Redmond, S., Sita, J. & Vincs, K. (2015)' Our Sherlockian Eyes: The Surveillance of Vision', *Refractory: Journal of Entertainment Media* 25. Retrieved from
http://refractory.unimelb.edu.au/2015/02/07/redmondsitavincs

Redmond, S. & Batty, C. (2015) 'Seeing into Things: Eye Tracking the Moving Image', *Refractory: Journal of Entertainment Media* 25. Retrieved from
http://refractory.unimelb.edu.au/2015/02/07/redmondbatty/

Saldanha G. & O'Brien S. (2013) *Research Methodologies in Translation Studies*, New York: Routledge.

Schaffer, R. (1982) 'Are There Consistent Vocal Cues for Irony?', in: C. S. Masek, A. Roberta & M. F. Miller (eds) *Papers from the Parasession on Language and Behaviour,* Chicago: Chicago Linguistic Society, 204–210.

Scharrer, L., Christmann, U. Knoll, M. (2011) ,Voice Modulations in German Ironic Speech', *Language and Speech* 54 (4): 435–465. Retrieved from
http://las.sagepub.com/content/54/4/435.full.pdf+html?hwshib2=authn%3A1420670267%
3A20150106%253A8f0f8e88-e38c-4fed-bcd9-
f811ef2601a4%3A0%3A0%3A0%3AjhlNsK%2FIsXsNXuB67DHheg%3D%3D

Schauffler, S. (2012) *Investigating Subtitling Strategies for the Translation of Wordplay in Wallace and Gromit – An Audience Reception Study* (Doctoral dissertation). University of Birmingham, Birmingham, UK.

Sepielak, K. (2016) 'Synchronization Techniques in Multilingual Fiction: Voiced-Over Films in Poland', *International Journal of Communication* (10): 1054–1073. Retrieved from
http://ijoc.org/index.php/ijoc/article/viewFile/3559/1578

Sepielak, K. & Matamala, A. (2014) 'Synchrony in the Voice-over of Polish Fiction Genres', *Babel* 62 (2): 145–163. Retrieved from
https://benjamins.com/#catalog/journals/babel.60.2.02sep/details

Schilperoord, J., de Groot, V. & van Son, N. (2005) 'Nonverbatim Captioning in Dutch Television Programs: A Text Linguistic Approach', *Journal of Deaf Studies and Deaf Education* 10 (4): 402−416. Retrieved from
https://academic.oup.com/jdsde/article/10/4/402/363399

Schmidt–Weigand, F., Kohnert, A. & Glowalla, U. (2010) 'A Closer Look at Split Visual Attention in System and Self-paced Instruction in Multimedia Learning', *Learning and Instruction* (20): 100–110. Retrieved from https://link.springer.com/article/10.1007/s10212-013-0171-8

Scollon, R., Scollon, S., W. (2014) 'Multimodality and lLanguage: A Retrospective and Prospective View', in: J. Carey (ed.) *The Routledge Handbook of Multimodal Analysis*, London: Routledge, 205–217.

Scott, A. O. (2009) 'The Brawling Supersleuth of 221B Baker Street Socks It to Em'. Rerieved from http://www.nytimes.com/2009/12/25/movies/25sherlock.html

Searle, J.R. (1969) *Speech Acts: An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Seto, K. I. (1998) 'On Non-echoic Irony', in: R. Carston & S. Uchida (eds) *Relevance theory: Applications and implications,* Amsterdama: John Benjamins Publishing, 239–255.

Shadrina, K. (2014) 'Social Hero Evolution in Mass Universal Cinema XX-XXI Centuries', *International Journal of Research in Social Sciences* 3 (8): 234–254. Retrieved from http://ijsk.org/uploads/3/1/1/7/3117743/6_sociology.pdf

Sita, J., Batty, C. & Perkins, C. (2015) 'How We Came To Eye Tracking Animation: A Cross-Disciplinary Approach to Researching the Moving Image', *Refractory: Journal of Entertainment Media* 25. Retrieved from http://refractory.unimelb.edu.au/2015/02/07/sitabattyperkins/

Smiley, S. (1971) *Playwriting: The Structure of Action*, New York: Prentice Hall.

Smith, T., J. & Mital P., K. (2013) 'Attentional Synchrony and the Influence of Viewing Task on Gaze Behaviour in Static and Dynamic Scenes', *Journal of Vision* 13 (8): 1–24. Retrieved from http://www.journalofvision.org/content/13/8/16

Smith, T., J. (2013) 'Watching You Watch Movies: Using Eye Tracking to Inform Cognitive Film Theory', in: A. P. Shimamura (ed.) *Psychocinematics: Exploring Cognition at the Movies,* New York: Oxford University Press, 165–191.

Smith, T., J. (2015) 'Read, Watch, Listen: A Commentary on Eye Tracking and Moving Images', *Journal of Entertainment Media*: *Eye-Tracking the Moving Image* 25. Retrieved from http://refractory.unimelb.edu.au/2015/02/07/smith/

Sowa, B. (2008) 'TVP odpowiada: Będą napisy zamiast lektora'. Retrieved from http://wiadomosci.dziennik.pl/wydarzenia/artykuly/127212,tvp-odpowiada-beda-napisy-zamiast-lektora.html

Specker, E., A. (2008) L1/L2 Eye Movement Reading of Closed Captioning: A multimodal analysis of multimodal use). A dissertation submitted to the University of Arizona for the degree of PhD.

Sperber, D. & Wilson, D. (1981) 'Irony and the Use–mention Distinction', in: P. Cole (ed.) *Radical Pragmatics,* New York: Academic Press, 295–318.

Sperber, D. & Wilson, D. (1986) *Relevance: Communication and Cognition*, Blackwell: Oxford.

Sperber, D. & Wilson, D. (1992) 'On Verbal Irony', *Lingua* (87): 53–76. Retrieved from http://www.ucl.ac.uk/ls/studypacks/Wilson-Onverbalirony.pdf

Sperber, D. & Wilson, D. (1995) *Relevance: Communication and Cognition*, Oxford: Blackwell.

Sperber, D. & Wilson, D. (1998) 'Irony and Relevance: A reply to Seto, Hamamoto and Yamanashi', in: R. Carston & S. Uchida (eds) *Relevance theory: Applications and implications,* Amsterdam: John Benjamins Publishing, 283–293.

Sperber, D. & Wilson, D. (2012) 'Explaining Irony', *Meaning and Relevance* 123–145. Retrieved from http://www.dan.sperber.fr/wpcontent/uploads/WilsonSperber_ExplainingIrony.pdf

Stafford, R. (2007) *Understanding Audiences and the Film Industry*, London: British Film Institute.

Strauss, A. L. & Corbin, J. M. (1990) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Thousand Oaks: Sage Publications.

Subbotko, D. (2008) 'Kapitulacja BBC Prima' [Capitulation of BBC Prime]. Retrieved from www.wyborcza.pl/1,76842,4827095.html

Sutton, T.M. & Altarriba, J. (2008) 'Representing Emotion Words in the Mental Lexicon: A New Look at the Emotional Stroop Effect', *The Mental Lexicon* (3): 19–46. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/1367006910379261

Szarkowska, A. (2009) 'The Audiovisual Landscape in Poland at the Dawn of the 21st Century', in: A. Goldstein & B. Gulobović (eds) *Foreign Language Movies: Dubbing vs. Subtitling.* Hamburg. Kovać, 185–201.

Szarkowska, A. (2011) 'Manipulation in Audiovisual Translation: At the crossroads of language, culture and politics', in: J. Esquibel & M. Kizeweter (eds) *Manipulation in translation,* Warszawa: Wydawnictwo SWPS, 158–165.

Szarkowska, A. & Laskowska, M. (2014) 'Jakie powinny być napisy? Raport z badania preferencji widzów na temat napisów telewizyjnych'. Retrieved from http://avt.ils.uw.edu.pl/files/2014/09/Wyniki-ankiety-o-napisach_Polska_FINAL.pdf

Szarkowska, A., Krejtz, I., Dutka, L., Pilipczuk, O. & Kruger J. L. (2016) 'The Effects of Text Editing and Subtitle Presentation Rate on the Comprehension and Reading Patterns of Interlingual and Intralingual Subtitles among Deaf, Hard of Hearing and Hearing Viewers', *Across Languages and Cultures* 17 (2): 183–204. Retrieved from http://akademiai.com/doi/pdf/10.1556/084.2016.17.2.3

Tallon, P. & Baggett, D. (2012) *The Philosophy of Sherlock Holmes,* Lexington: University Press of Kentucky.

Tan, S., E. (1995) *Emotion and the Structure of Film Narrative: Film as an Emotion Machine*, New York: Routledge.

Taylor, Ch. (2003) 'Multimodal Transcription', *The Translator* 9 (2): 191–205.

Thomas, A. (2006) *Value and Context: The Nature of Moral and Political Knowledge,* Oxford: Oxford University Press.

Tomaszkiewicz, T. (2006) *Przeklad Audiovisualny,* Warszawa: Wydawnictwo Naukowe PWN.

Treuting, J. (2006) 'Eye Tracking and Cinema: A Study of Film Theory and Visual Perception', *Society of Motion Picture and Television Engineers* 115 (1): 31–40. Retrieved from http://ieeexplore.ieee.org/document/7264815/

Turner, K., L. (2016) *This is the Sound of Irony: Music, Politics and Popular Culture,* New York: Routledge.

Unsworth, L. & Cleirigh, Ch. (2009) 'Multimodality and Reading: The Construction of Meaning through Image-text Interaction', in: J. Carey (ed.) *The Routledge Handbook of Multimodal Analysis*, London: Routledge, 151–164.

Utsumi, A. (2000) 'Verbal Irony as Implicit Display of Ironic Environment: Distinguishing Ironic Utterances from Nonirony', *Journal of Pragmatics* (32): 1777–1806. Retrieved from http://ac.els-cdn.com/S0378216699001162/1-s2.0-S0378216699001162-main.pdf?_tid=b3a3cc30-3153-11e4-bcb0-00000aab0f6b&acdnat=1409519809_f79dbbd6a4bf55d7f488a5bbff713a3a

Webb, N. & Renshaw, T. (2008) 'Eye Tracking HCI', in: P. Cairns & A. Cox (eds) *Research Methods for Human-Computer Interaction*, Cambridge: Cambridge University Press.

Wharton, D. & Grant, J. (2005) *Teaching Analysis of Film Language*, London: British Film Institute.

Wierzbicka, A. (2003) *Cross-cultural Pragmatics: The Semantics of Human Interaction,* Berlin: Walter de Gruyter.

Wierzbicka, A. (2004) 'Happiness in Cross-linguistic and Cross-cultural Perspective, *Daedalus* 133 (2): 34–43. Retrieved from https://www.researchgate.net/publication/240677172_'Happiness_in_cross-linguistic_cross-cultural_perspective

Williams, J.A., Burns, E.L. & Harmon, E.A. (2009) 'Insincere Utterances and Gaze: Eye Contact during Sarcastic Statements', *Perceptual and Motor Skills* 08 (2): 565–572.

Wilson, D. (2006) 'The Pragmatics of Verbal Irony: Echo or Pretense?', *Lingua* (116): 1722–1743. Retrieved from http://www.sciencedirect.com/science/article/pii/S0024384106001124#

Wilson, D. (2013) 'Irony Comprehension: A developmental Perspective', *A Journal of Pragmatic*s (59): 40–56. Retrieved from https://www.duo.uio.no/bitstream/handle/10852/58988/Wilson%2B2013%2BIrony%2Bcomprehension%2B-%2Ba%2Bdevelopmental%2Bperspective.pdf?sequence=1&isAllowed=y

Wilson, D. (2012) 'Sherlock Holmes and The Case of The Edinburgh Haunting'. Retrieved from www.eighbooks.com/lib.php?q=sherlock-holmes-and-the-case-of-the-edinburgh-haunting

Woźniak, M. (2012) 'Voice-over or Voice-in-between? Some Considerations about the Voice-over Translation of Feature Films on Polish Television', *Approaches to Translation Studies* (36): 209–229. Retrieved from www.academia.edu/1882232/Voice-over_or_Voice-in-between_Some_Considerations_about_the_Voice-over_Translation_of_Feature_Films_on_Polish_Television

van Leeuwen, T. (2014) 'Parametric Systems: The case of voice quality', in: J. Carey (ed.) *The Routledge Handbook of Multimodal Analysis,* London: Routledge, 76–86.

van Leeuwen, T. (2014), 'Colour: Code, Mode, Modality', in: J. Carey (ed.) *The Routledge Handbook of Multimodal Analysis,* London: Routledge, 397–410.

van Leeuwen, T. (1999) *Speech, Music, Sound*, London: Palgrave Macmillan.

Vanacker, S. & Wynne, C. (2013) *Sherlock Holmes and Conan Doyle: Multi-Media Afterlives,* New York: Palgrave Macmillan.

Vandaele, J. (1999) 'Each Time We Laugh. Translated Humour in Screen Comedy', in: J. Vandaele (ed.) *Translation and the (Re)Location of Meaning. Selected Papers of the CETRA Research Seminars in Translation Studies,* Leuven: CETRA, 237–272.

Vanoye, F. (1985) 'Conversations publiques', *Iris* 3 (1): 99–188. Google Scholar.

Vinay, J., P. & Darbelnet, J. (1995) *Comparative Stylistics of French and English: A Methodology for Translation*, Amsterdam: John Benjamins Publishing.

Zabalbeascoa, P. (2003) 'Translating Audiovisual Screen Irony', in: L. Pérez-González (ed) *Speaking in Tongues: Languages across contexts and Users* Valencia: University of Valencia, 305–322. Retrieved from http://www.academia.edu/3239334/Translating_Audiovisual_Screen_Irony

Zabalbeascoa, P. (2008) 'The Nature of the Audiovisual Text and its Parameters', in: J. Diaz Cintas (ed) *The Didactics of Audiovisual Translation*, Amsterdam: John Benjamins Publishing, 21–37.

Yus Ramos, F. (1997-98) 'Irony: Context Accessibility and Processing Effort', *Pragmalingüística* 5 (6): 391– 410. Retrieved from http://rodin.uca.es/xmlui/bitstream/handle/10498/8781/17218457.pdf?sequence=1

Yus, F. (2000) 'On Reaching the Intended Ironic Interpretation', *International Journal of Communication* 10 (1–2): 27–78. Retrieved from https://www.academia.edu/2013213/On_reaching_the_intended_ironic_interpretation

Electronic Resources

Online dictionary PWN of Polish language, http://sjp.pwn.pl

Dziennik Gazeta Prawna, Available at: http://wiadomosci.dziennik.pl/wydarzenia/artykuly/127212,tvp-odpowiada-beda-napisy-zamiast-lektora.html (Accessed: 22/10/2016)

ADD Statistical Yearbook in film industry. [Online] Available at: http://www.bfi.org.uk/education-research/film-industry-statistics-research/statistical-yearbook (Accessed: 05/11/2015)

Filmography & TV Series

Sherlock Holmes (2009), Guy Ritchie, UK and USA.

Sherlock Holmes: A Game of Shadows (2011), Guy Ritchie, UK and USA.

Transpotting (199), Danny Boyle, USA.

House, M.D. (2004-2012), Omar Epps, USA.

The Big Boss (1971) Lo Wei, China and USA.

Appendices

Appendix 1 - Consent form, information sheet and CEFL framework

MANCHESTER 1824

The University of Manchester

Investigation the Reception of On-screen Information in Film Translation –

An Experimental Approach

CONSENT FORM

Jeśli zgadasz się wziąć udział w badaniu proszę wypełnij I podpisz formularz

Proszę wpisz parafkę

1. Rozumiem, że mój udział w badaniu jest całkowicie dobrowolny i mogę wycofać się z niego w każdej chwili bez podania przyczyny i bez szkoday dla prowadzonego badania.	
2. Rozumiem, że ruchy moich oczu będą nagrywane technologią eye-trackingową.	
3. Zgadzam się do użycia anonimowych cytatów	
4. Zgadzam się wypełnić ankietę dotyczącą scen wykorzytywanych w badaniu.	

Zgadzam się na udział w wyżej wymienionym projekcie

_____   _____   _____
Imię i nazwisko uczestnika   Data   Podpis

_____   _____   _____
Imię   i   nazwisko   osoby   Data   Podpis
wykonującej badanie

_____   _____   _____

Investigation the Reception of On-screen Information in Film Translation –

An Experimental Approach

Participant Information Sheet

Jesteś zaproszony do udziału w badaniu naukowym które ma na celu sprawdzenie jak widzowie odbierają informację przekazana w wybranych scenach filmów *Sherlock Holmes* (2009) i *Shelock Holmes: A Game of Shadows* (2011). Badanie jest częśćią projektu doktoranckiego. Zanim zdecydujesz się na udział ważne jest abyś zrozumiał/ła dlaczego eksperyment jest wykonywany i co.... Proszę przeczytaj uważnie informacie poniżej i zapytaj jeśli coś nie jest jasne lub jeśli chciałbyś/łabys otrzymać więcej informacji. Zastanów się czy chcesz uczestniczyć czy nie. Dziękuję za przeczytanie tej informacji.

Kto będzie przeprowadzał badanie?

Paulina Burczynska, Student doktorancki

Centre for Translation and intercultural Studies
School of Languages, Arts and Cultures
The University of Manchester
Oxford Road,
Manchester,
M13 9PL,
UK

Tytuł projektu

Investigation the Reception of On-screen Information in Film Translation –

An Experimental Approach

Moje badanie ma głównie na celu ustalenie jak polscy widzowie procesują informację pokazaną na ekranie podczas oglądania wybranych scen z filmu Sherlock Holmes (2009) and Sherlock Holmes: A Game of Shadows (2011) tłumaczonych dwoma różnymi technikami tłumaczeniowymi, tzn. Z napisami z i lekotrem.

Jaki jest cel badania?

Mam nadzieję, ze uda mi się dowiedzieć jak widzowie rozumieją informację przekazaną w dwóch tłumaczeniach filmów wybranych do eskperymentu.

228

Dlaczego zostałem/-łam zaproszony/a do badania?

Będziesz jednym z 30 uczestników, którzy zostali poproszeni o udział w badaniu. Twoje zgłoszeni zostało zaakceptowane, ponieważ spełniach dwa kryteria: jesteś narodowości polskiej oraz jesteś atudentem lub absolwneten w wieku 19-35 lat.

O co będę poproszony/a jeśli zdecyduję się na udział?

First, you will be asked to complete the pre-experiment questionnaire that will collect the information regarding your age, sex, English language proficiency and familiarity with the audiovisual material used in the present study. Second, you will be asked to watch 4 scenes, and then you will be asked to answer several questions about them. The whole experiment will take around 90 min. After 45 min, you will be offered an opportunity to take a break. You will be asked to seat in a natural in from of the screen as your eye movements will be recorded with the eye-tracking technology.

Co się stanie z zabranymi danymi?

Zabrane dane bedą znalalizowane i wykorzystane na potrzeby projekktu doktoranckiego.Wyniki mogą różnież być opublikowane w formie ...artukułów lub zaprezentowane na międzynanrodowej konferencji. Date będą znanalizowane i opublikowane anonimowo.

How is confidentiality maintained?

The data will be stored by the university and on the encrypted researcher's laptop for the period of 5 years. Afterwards the data will be destroyed.

Co się stanie jeśli nie będę chciał/ła uczestniczyć lub zmienię zdanie?

Decyzja o uczestniczeniu w badaniu należy do Ciebie. Jeśli zdecydujesz wziąć udział, będziesz poproszony o zachowanie tego formularza i o podpisanie zgody na badanie. Jeśli zdecydujesz się wziąć udział wciąż możesz się wycofać bez podania przyczyny i bez szkody na swojej osoby.

Czy dostanę wynagrodzenie pieniężne za udział w badaniu?

Ponieważ jest to udział dobrowolny, uczestnicy nie będa otrzymywać wynagrodzenia pieniężnego za udział w eksperymencie. Jednakże, koszty dojazdów do i z laboratorum gdzie badanie będzie się odbywać będą zwracane uczestnikom.

Jak długo trwa badanie?

Całe badanie nie będzie trwać dłuzej niż 90 minut włączając w to czas na obejrzenie scen, odpowiedzenie na ankietę do scen i ankietę demograficzną.Będziesz zaoferowany przerwę po 45 minutach.

Gdzie będzię się odbywało badanie?

Eksperyment będzie się odbywał na Uniwersytecie Warszawskim:
Instytut Neofilologii

Ul. Dobra 55
Sala 0.096 (na parterze, wchodząc głownym wejsciem skręcić w lewo i iść do konca korytarza)
Warszawa

Gdzie będą publikowane wyniki badania?

Wyniki badania będą publikowane w formie ...artuykułów lub prezentowane na międzynarodowych konferencjach.

Kto zaakceptował ten projekt?

Projekt został ...I zaakceptowany przez Komisję Etyczną Uniwersytetu w Manchesterze dnia 27/01/2015.

Contact for further information

Doktorant wykonujący badanie:

Paulina Burczynska, Paulina.burczynska@postgrad.manchester.ac.uk

Promotor Dr Luis Perez-Gonzalez luis.perez-gonzalez@manchester.ac.uk

A jeśli coś pójdzie źle?

Jeśli uważasz, że istnieją aspekty badania o których nie chcesz rozmawiać z członkami zespołu badawczego, skontaktuj się z koordynatorem Praktyk Badawczych i Governance pisenie lub meilowo:

Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL', by emailing: Research-Governance@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.

Self-assessment of foreign language competences (CEFL)

Kompetentne posługiwanie się językiem obcym		
	C2	Potrafi zrozumieć bez wysiłku praktycznie wszystko, co czyta lub słyszy. Potrafi odtworzyć fakty i argumenty z różnych źródeł pisemnych i ustnych streszczając je w sposób zwięzły i spójny. Potrafi wypowiedzieć się spontanicznie, bardzo płynnie i precyzyjnie oraz uwydatnić niuanse znaczeniowe tekstów o złożonej tematyce.
	C1	Potrafi zrozumieć szeroką gamę długich i trudnych tekstów oraz zrozumieć ukryte w nich podteksty. Potrafi wypowiedzieć się spontanicznie i biegle nie zastanawiając się zbytnio nad doborem słów. Potrafi posługiwać się językiem skutecznie i swobodnie w życiu społecznym, zawodowym lub w czasie studiów. Potrafi budować wypowiedzi na tematy złożone, jasne, o wyraźnej strukturze i wykazać się opanowaniem narzędzi językowych służących organizacji i wewnętrznej spójności dyskursu.

Samodzielne posługiwanie się językiem obcym	B2	Potrafi zrozumieć zasadnicze aspekty problemów konkretnych lub abstrakcyjnych przedstawionych w tekstach złożonych, w tym dyskusję specjalistyczną dotyczącą własnej tematyki zawodowej. Potrafi porozumieć się na tyle swobodnie i spontanicznie, że rozmowa z rdzennym użytkownikiem języka wolna jest od napięć, tak w przypadku jednej jak i drugiej strony. Potrafi wyrazić się w sposób jasny i szczegółowy na wiele tematów, wyrazić opinię na dany temat wykazując pozytywne i negatywne strony różnych (proponowanych) wyborów.
	B1	Potrafi zrozumieć zasadnicze punkty rozmowy, gdy używany jest język jasny i standardowy, a rozmowa dotyczy spraw znanych związanych z pracą, szkołą, czasem wolnym, etc. Potrafi sobie poradzić w większości sytuacji, jakie spotyka w podróży w regionie języka docelowego. Potrafi wypowiedzieć się w sposób prosty i zwięzły na tematy z życia codziennego i dotyczące własnych zainteresowań. Potrafi opowiedzieć wydarzenie, przeżycie osobiste lub sen, wyrazić nadzieję lub cel, jak również przedstawić krótko uzasadnienie lub wyjaśnienie dotyczące projektu lub pomysłu.
Podstawowe posługiwanie się językiem obcym	A2	Potrafi zrozumieć pojedyncze zdania oraz wyrażenia często używane i związane bezpośrednio z życiem codziennym (np.: dane o sytuacji osobistej i rodzinnej, zakupy, najbliższe otoczenie, praca). Potrafi się porozumieć w trakcie wykonywania zadań prostych, codziennych, wymagających jedynie bezpośredniej i prostej wymiany informacji na znany temat. Potrafi opisać za pomocą prostych środków swoje wykształcenie, swoje bezpośrednie otoczenie i wypowiadać się na tematy związane z niezbędnymi potrzebami.
	A1	Potrafi zrozumieć i stosować wyrażenia z języka codziennego oraz wyrażenia potoczne, jak również bardzo proste wypowiedzi mające na celu zaspokojenie konkretnych potrzeb. Potrafi przedstawić siebie lub inną osobę oraz zadać jej pytania dotyczące – np. miejsca zamieszkania, jej znajomych, posiadanych przez nią rzeczy, etc. – może także odpowiedzieć na pytania tego samego typu. Potrafi porozumiewać się z innymi w prosty sposób, jeśli jego rozmówcy mówią wolno i wyraźnie oraz wykazują chęć współpracy.

Appendix 2 - Pre-experimental questionnaire

Nr pytania (BT: Question No)	Znajomość filmu i charakteru Sherlocka Holmsa (BT: Familiarity with the film and the character of Sherlock Holmes)	Odpowiedź (BT: Response)
1	Czy posiadasz ogólne informacje o postaci Sherlocka Holmesa? (np. z filmów, książek, artykułów, seriali telewizyjnych) (BT: Have you got any information about the character of Sherlock Holmes? E.g. from films, books, articles, TV series)	Pytanie otwarte (BT: Open-ended question)
2	Czy oglądałeś te wersje filmowe Sherlocka Holmesa w reżyserii Guya Ritchie? (BT: Did you watch the versions of Sherlock Holmes directed by Guy Ritchie?)	Tak - *Sherlock Holmes* (BT: Yes - *Sherlock Holmes)* Nie - *Sherlock Holmes* (BT: No - *Sherlock Holmes*)
3	Czy oglądałeś te wersje filmowe Sherlocka Holmesa w reżyserii Guya Ritchie? (BT: Did you watch the versions of Sherlock Holmes directed by Guy Ritchie?)	Tak - *Sherlock Holmes: Gra Cieni* (BT: Yes - *Sherlock Holmes : A Game of Shadows*) Nie - *Sherlock Holmes: Gra Cieni* (BT: No - *Sherlock Holmes: A Game of Shadows*)

Nr pytania (BT: Question No)	Ankieta samo oceniająca biegłość w j. angielskim i upodobania audiowizualne. (BT: Self-assessment questionnaire about proficiency in English and audiovisual habits)	Odpowiedź (BT: Response)
1	Jak długo uczysz się języka angielskiego? (BT: How long have you been learning English?)	Nigdy nie uczyłam (-łem) się języka angielskiego: (BT: I have never learnt English) 1-2 lata (BT: 1-2 years) 3-4 lata (BT: 3-4 years) 4-6 lat (BT: 4-6 years) 7-10 lat (BT: 7-10 years) 10 lat i więcej (BT: 10 years and more)
2	Ile czasu spędziłaś (-łeś) w anglojęzycznym kraju lub otoczeniu? (BT: How much time have you spent in an English-speaking country or environment?)	W ogóle (BT: None) Kilka dni (BT: A few days) Kilka tygodni (BT: A few weeks) kilka miesięcy (BT: A few months) 1 rok i więcej (BT: 1 year and more)
3	Jak oceniasz swoją umiejętność słuchania i rozumienia w języku angielskim? (BT: How do you assess your listening comprehension in English)	A1 - Początkujący (BT: A1 - Beginner) A2 - Niższy średnio-zaawansowany (BT: A2 - Lower intermediate) B1 - Średnio-zaawansowany (BT: B1 - Intermediate) B2 - Wyższy średnio-zaawansowany (BT: B2 - Upper intermediate) C1 - Zaawansowany (BT: C1 - Advanced) C2 - wyższy zaawansowany (BT: C2 - Upper advanced)
4	Którą metodę tłumaczenia preferujesz oglądając filmy w języku angielskim? (BT: What translation method you prefer when you watch films in English?)	Napisy (BT: Subtitles) Lektor (BT: Voice-over)
5	Jak często oglądasz filmy w języku angielskim z lektorem? (BT: How often do you watch films in English with voice-over?)	Codziennie (BT: Everyday) Kilka razy w tygodniu (BT: A few times per week) Kilka razy w miesiącu (BT: A few times per month) Kilka razy w roku (BT: A few times per year) Nie oglądam w ogóle (BT: I do not watch at all)
6	Jak często oglądasz filmy w języku angielskim z napisami? (BT: How often do you watch films in English with subtitles?)	Codziennie (BT: Everyday) Kilka razy w tygodniu (BT: A few times per week) Kilka razy w miesiącu (BT: A few times per month) Kilka razy w roku (BT: A few times per year) Nie oglądam w ogóle (BT: I do not watch at all)

Appendix 3 - Experimental Questionnaire

|	Pytania do każdego klipu (BT: Questions to each clip)	Odpowiedzi (BT: Responses)
SH1_1	*Cmentarz i Lord Blackwood* (BT: *Cemetery and Lord Blackwood*)	
1	Dlaczego Sherlock mówi do Inspektora Lestrada: "Na jakim etapie? Zastanawiania się?" (BT: Why is Sherlock responding to Inspector Lestrade: "At what stage of the process? Contemplative?")	Pytanie otwarte (BT: Open-ended question)
2	Jak byś opisał(a) stosunek Sherlocka do Lestrada i policjantów z Scotland Yardu? (BT: How would you describe Sherlock's attitude towards Lestrade and the Scotland Yard policemen?)	Pytanie otwarte (BT: Open-ended question)
3	Co dokładnie (np. gesty, ruchy ciała, dźwięki) pomogło Ci opisać relację Sherlocka z Lestradem i policjantami z Scotland Yardu w ten sposób? (BT: What exactly (e.g., gestures, body movements, sounds) helped you describe the relation between Sherlock and Lestrade with the Scotland Yard policemen in this way?	Pytanie otwarte (BT: Open-ended question)
4	Czy sądzisz, że tłumaczenie z napisami utrudniło Ci w jakimś stopniu zrozumieć to co Sherlock chciał przekazać w tym fragmencie? Jeśli wybrałeś odpowiedź 1 lub 2 proszę uzasadnij w polu poniżej. (BT: Do you believe that subtitles hindered your comprehension to some extent of what Sherlock intended to relay in the fragment? If you selected the response 1 or 2, please justify in the box below.)	1 Tak (BT: Yes) 2 Raczej tak (BT: Rather yes) 3 Nie wiem (BT: I don't know) 4 Raczej nie (BT: Rather no) 5 Nie (BT: No)
SH1_2	*Watson i Sherlock* (BT: *Watson and Sherlock*)	
1	Dlaczego Watson zwraca się do Sherlocka mówiąc: "Wyglądasz cudownie"? (BT: Why is Watson responding to Sherlock: "You look gorgeous"?)	Pytanie otwarte (BT: Open-ended question)
2	Jaki byś określił(a) sposób w jaki Watson zwraca się do Sherlocka w tej scenie? (BT: How would you describe the way in which Watson is talking to Sherlock in this scene?)	Pytanie otwarte (BT: Open-ended question)
3	Czy zauważyłeś (-łaś) lub usłyszałeś (-łaś) podczas tej sceny elementy takie jak gesty, ruchy ciała, scenerię które ułatwiły Ci zrozumieć to co Watson miał na myśli? (BT: Have you seen or heard any elements in this scene such as gestures, body movements, settings that helped you understand what Watson meant?)	Pytanie otwarte (BT: Open-ended question)
4	Czy myślisz, że tłumaczenie z napisami zakłóciło Ci w jakiś sposób zrozumieć to co Watson miał na myśli w tej scenie? Jeśli wybrałeś odpowiedź 1 lub 2 proszę uzasadnij w polu poniżej. (BT: Do you believe that subtitles hindered your comprehension to some extent of what Watson meant in this scene? If you selected the response 1 or 2, please justify in the box below.)	1 Tak (BT: Yes) 2 Raczej tak (BT: Rather yes) 3 Nie wiem (BT: I don't know) 4 Raczej nie (BT: Rather no) 5 Nie (BT: No)

	Pytania do każdego klipu (BT: Questions to each clip)	**Odpowiedzi (BT: Responses)**
SH1_3	*Sherlock i Lord Coward* (BT: *Sherlock and Lord Coward*)	
1	Dlaczego Sherlock zwraca się do Lorda Cowarda mówiąc: "A więc nie ma czasu do stracenia, prawda?" (BT: Why does Sherlock respond to Lord Coward: "There isn't any time to waste then. Is there?")	Pytanie otwarte (BT: Open-ended question)
2	Jak myślisz jaki jest stosunek Sherlocka do Lorda Cowarda w tym fragmencie? (BT: Do you think what is Sherlock's attitude towards Lord Coward in this fragment?)	Pytanie otwarte (BT: Open-ended question)
3	Jakie elementy ułatwiły Ci zinterpretować to co Sherlock chciał przekazać wypowiadając powyższe zdanie? Np. gesty, ruchy ciała, dźwięki (BT: What elements helped you interpret what Sherlock intended to relay in the dialogue above? E.g. gestures, body movements, sounds)	Pytanie otwarte (BT: Open-ended question)
4	Czy uważasz, że tłumaczenie z napisami przeszkodziło Ci w jakimś stopniu zrozumieć to co Sherlock chciał przekazać w tym momencie? Jeśli wybrałeś odpowiedź 1 lub 2 proszę uzasadnij w polu poniżej. (BT: Do you believe that subtitles hindered your comprehension to some extent of what Sherlock intended to relay in the fragment? If you selected the response 1 or 2, please justify in the box below.)	1 Tak (BT: Yes) 2 Raczej tak (BT: Rather yes) 3 Nie wiem (BT: I don't know) 4 Raczej nie (BT: Rather no) 5 Nie (BT: No)
SH2_1	*Sherlock, Watson i szalik* (BT: *Sherlock, Watson and a scarf*)	
1	Dlaczego Sherlock zwraca się do Watsona mówiąc: "Wolisz schludny strój wojskowego w połączeniu z szpetnym szalikiem wydzierganym przez narzeczoną?" (BT: Why is Sherlock responding to Watson: "Do you prefer I joined you in the fashion faux pas of wearing fine military dress with that heinous handmade scarf clearly one of your fiancée's early efforts?")	Pytanie otwarte (BT: Open-ended question)
2	Jak byś opisał(a) sposób w jaki Sherlock odnosi się do Watsona w tym fragmencie? (BT: How would you describe the way in which Sherlock is talking to Watson in this fragment?)	Pytanie otwarte (BT: Open-ended question)
3	Co pomogło Ci zinterpretować stosunek Sherkocka do Watsona w tych słowach? (np. gesty, ruchy, dźwięki, sceneria) (BT: What helped you interpret Sherlock's attitude towards Watson in this way?)	Pytanie otwarte (BT: Open-ended question)
4	Czy sądzisz, że tłumaczenie z lektorem utrudniło Ci w jakiś sposób zrozumieć to co Sherlock chciał wyrazić w tej scenie? (Jeśli wybrałeś odpowiedź 1 lub 2 proszę uzasadnij w polu poniżej) (BT: Do you believe that voice-over hindered your comprehension to some extent of what Sherlock intended to relay in the fragment? If you selected the response 1 or 2, please justify in the box below.)	1 Tak (BT: Yes) 2 Raczej tak (BT: Rather yes) 3 Nie wiem (BT: I don't know) 4 Raczej nie (BT: Rather no) 5 Nie (BT: No)

	Pytania do każdego klipu (BT: Questions to each clip)	Odpowiedzi (BT: Responses)
SH2_2	*Wieczór kawalerski Watsona* (BT: *Watson's stag party*)	
1	Dlaczego Mycroft zwraca się do Watsona mówiąc: "Szkoda, że przyjaciele nie przyszli". (BT: Why is Mycroft saying to Watson: "Shame none of your friends could make it".)	Pytanie otwarte (BT: Open-ended question)
2	Jak byś określił(a) sposób w jaki Mycroft zwraca się do Watsona w tej scenie? (BT: How would you describe the way in which Mycroft is talking to Watson in this scene?)	Pytanie otwarte (BT: Open-ended question)
3	Czy zauważyłeś (-łaś) lub usłyszałeś (-łaś) podczas tej sceny elementy takie jak gesty, ruchy ciała, dźwięki, scenerię, które ułatwiły Ci zrozumieć co Mycroft chciał przekazać? (BT: Have you seen or heard any elements in this scene such as gestures, body movements, sounds, scenery that helped you understand what Mycroft meant to say?)	Pytanie otwarte (BT: Open-ended question)
4	Czy myślisz, że tłumaczenie z lektorem zakłóciło Ci w jakimś stopniu zrozumieć to co Mycroft chciał przekazać w tym momencie? Jeśli wybrałeś odpowiedź 1 lub 2 proszę uzasadnij w polu poniżej. (BT: Do you believe that voice-over hindered your comprehension to some extent of what Mycroft intended to relay in the fragment? If you selected the response 1 or 2, please justify in the box below.)	1 Tak (BT: Yes) 2 Raczej tak (BT: Rather yes) 3 Nie wiem (BT: I don't know) 4 Raczej nie (BT: Rather no) 5 Nie (BT: No)
SH2_3	*Zasadzka w pociągu* (BT: *The ambush on a train*)	
1	Dlaczego Sherlock mówi do Watsona: "Mówiłem, celnie! Ile szans mam ci dawać?" (BT: Why is Sherlock responding to Watson: "I said make it count. How many windows must I provide?")	Pytanie otwarte (BT: Open-ended question)
2	Jak byś opisał(a) sposób w jaki Sherlock odnosi się do Watsona wypowiadając powyższe zdanie? (BT: How would you describe Sherlock's attitude towards Watson when uttering in the sentence above?)	Pytanie otwarte (BT: Open-ended question)
3	Jakie elementy (np. sceneria, dźwięki, ruchy) pomogły Ci zinterpretować to co Sherlock chciał przekazać w tym fragmencie? (BT: What elements (e.g., settings, sounds, movements) helped you interpret what Sherlock intended to relay in this fragment?)	Pytanie otwarte (BT: Open-ended question)
4	Czy myślisz, że tłumaczenie z lektorem przeszkodziło Ci w jakiś sposób zrozumieć to co Sherlock chciał wyrazić w tym fragmencie? (Jeśli wybrałeś odpowiedź 1 lub 2 proszę uzasadnij w polu poniżej) (BT: Do you believe that voice-over hindered your comprehension to some extent of what Sherlock intended to relay in the fragment? If you selected the response 1 or 2, please justify in the box below.)	1 Tak (BT: Yes) 2 Raczej tak (BT: Rather yes) 3 Nie wiem (BT: I don't know) 4 Raczej nie (BT: Rather no) 5 Nie (BT: No)

Appendix 4 - Post-experimental Questionnaire

Nr pytania (BT: Question No)	Elementy języka filmowego i pytania demograficzne (BT: Elements of the language of film and demographic questions)	Odpowiedzi (BT: Responses)
1	Jakie elementy języka filmowego z podanych poniżej pomogły Ci zinterpretować znaczenie wypowiedzi w obejrzanych scenach? (Język filmu to sposób organizowania materiałów wizualnych i dźwiękowych, zarejestrowanych na różnych nośnikach.) (BT: What elements of the language of film helped you interpret the meaning of the dialogues in the presented clips? The language of film is a way of organising visual and audio materials recorded on various media.)	Uniesione brwi (BT: Raised eyebrows) Obniżone brwi (BT: Lowered eyebrows) Oczy szeroko otwarte (BT: Wide-open eyes) Przewracanie oczami (BT: Rolling eyes) Mrugnięcie (BT: Winking) Lekceważące spojrzenie (BT: A disapproved look) Kamienna twarz (bez emocji) (BT: A stoned face) Kiwanie głową (BT: Nodding) Wzruszenie ramionami (BT: Shrugging) Złośliwy/pogardliwy uśmiech (BT: A smirk / contemptuous smile) Ruchy ciała (BT: Body movements) Śmiech (BT: Laughter) Muzyka instrumentalna (BT: Instrumental music) Wysoki ton głosu (BT: A high tone of voice) Umiarkowany ton głosu (BT: A medium tone of voice) Niski ton głosu (BT: A low tone of voice) Intonacja rosnąca (BT: Rising intonation) Intonacja malejąca (BT: Decreasing intonation) Intonacja płaska (BT: Flat intonation) Wolne tempo mówienia (BT: Slow speech rate) Umiarkowane tempo mówienia (BT: Moderate speech rate) Szybkie tempo mówienia (BT: Fast speech rate) Głośność mówienia - cicho (BT: Speaking volume - quiet) Głośność mówienia - normalnie (BT: Speaking volume - normal) Głośność mówienia - głośno (BT: Speaking volume - loud) Głośność mówienia - bardzo głośno (BT: Speaking volume - very loud) Przedłużone sylaby (BT: Extended syllables) Cisza / pauza (BT: Silence / Pause) Dźwięki para-lingwistyczne np. śmiech, płacz, ziewanie, westchnienia (BT: Paralinguistic sounds e.g., laughter, crying, yawning, sighing) Sceneria, rekwizyty, kostiumy (BT: Settings, props, costumes) Oświetlenie (BT: Lighting) Kąt (perspektywa) kamery (BT: Camera angle) Odległość kamery (BT: Camera distance) Pozycja kamery (BT: Camera position) Montaż (BT: Editing)
2	Ile masz lat? (BT: How old are you?)	
3	Co studiujesz lub studiowałaś (-łeś)? (BT: What do/did you study at university?)	

Appendix 5 - Multimodal Transcription
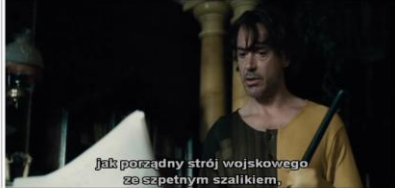
SH1_2

| Frame No	Time (hh:mm:ss)	Frame	Visual Image	Kinesic Action	Soundtrack	Subtitling	Voice-over
1	1:27:33		CP: stationary (vertical) CD: MLS, HA VC: settings, old room, sign of the floor, props, costumes ED: establishing shot VF: Sherlock wakes up. Watson looks at Sherlock from above.	⟷ Watson sits straight on the chair looking on Sherlock from above. Sherlock slowly sits on the bed.	. ♫	X	X
2	1:27:35		CP: stationary (horizontal) CD: MCS, SA VC: settings, old room, sign of the floor, props, costumes ED: establishing shot VF: Watson watched Sherlock as he wakes up.	⊙ ▭ (><) ⟷ Watson sits straight on a chair. He raises his head looking at Sherlock from above squinting his eyes. Watson smiles and Sherlock with pity.	♫ WAT: You look gorgeous. p, S, LP, # ↔	WAT Wyglądasz cudownie. [BT: You look gorgeous.]	WAT: Wyglądasz cudownie. [BT: You look gorgeous.]
3	1:27:36		CP: stationary (horizontal) CD: MCS, SA VC: settings, old room, sign of the floor, props, costumes ED: establishing shot VF: Watson watches Sherlock as he waked up.	△ ⊙ ▭ (><) ⟷ Watson sits straight on a chair. He raises his head looking at Sherlock from above squinting his eyes. Watson smiles ⊞	♫ WAT: You look gorgeous. p, S, LP, # ↔	WAT: Wyglądasz cudownie. [BT: You look gorgeous.]	WAT: Wyglądasz cudownie. [BT: You look gorgeous.]
4	1:27:37		CP: stationary (horizontal) CD: MCS, SA VC: settings, old room, sign of the floor, props, costumes ED: establishing shot VF: Watson watches Sherlock as he wakes up.	◇ ▭ ⟷ Watson moves his head and keeps it straight. and looks at Sherlock showing 'blank face'.	♫ WAT: You look gorgeous. p, S, LP, # ↔	WAT: Wyglądasz cudownie. [BT: You look gorgeous.]	WAT: Wyglądasz cudownie. [BT: You look gorgeous.]
5	1:27:38		CP: stationary (horizontal) CD: MCS, SA VC: settings, old room, sign of the floor, props, costumes ED: establishing shot VF: Watson watches Sherlock as he wakes up.	⊙ ⟷ Watson keeps his head straight. and looks at Sherlock smiling slightly.	♫ WAT: You look gorgeous. p, S, LP, # ↔	WAT: Wyglądasz cudownie. [BT: You look gorgeous.]	WAT: Wyglądasz cudownie. [BT: You look gorgeous.]
6	1:27:39		CP: stationary (horizontal) CD: MLS, SA VC: settings, old room, sign of the floor, props, costumes ED: establishing shot VF: Watson watches Sherlock as he wakes up.	⟷ Sherlock sit on the bed looking at the cloth that he holds in his hands. He avoid looking at Watson.	♫	X	X

SH1_3

| Frame No	Time (hh:mm:ss)	Frame	Visual Image	Kinesic Action	Soundtrack	Subtitles	Voice-over
1	1:38:38		CP: stationary (vertical) CD: MLS, SA VC: settings, smoke, LK ED: establishing shot VF: Sherlock looks at Coward standing next to him.	⟷ Sherlock sits comfortably on a chair behind Coward smoking is pipe while Coward is looking for him in the room.	♫	X	X
2	1:38:39		CP: stationary (horizontal) CD: CS, SA VC: LK, pipe, smoke ED: establishing shot VF: Sherlock observes Coward.	◎ △ ◇ ⊟ (><) ⟹ Sherlock sits calmly on a chair smoking his pipe. He tilts his head to right side, keeps his eyes wide open and his mouth firmly closed. expressing his 'blank face'.	♫ SH: *There isn't anytime to waste then.* **p, S, LP** ⟷	SH: A więc nie ma czasu do stracenia. [BT: So there is no time to lose.]	SH: A więc nie ma czasu do stracenia. [BT: So there is no time to lose.]
3	1:38:40		CP: stationary (horizontal) CD: CS, SA VC: LK, pipe, smoke ED: establishing shot VF: Sherlock observes Coward.	◎ △ ◇ ⊟ (><) ⟹ Sherlock sits calmly o na chair smoking his pipe. He tilts his head to right side, keeps his eyes wide open expressing his 'bank face'.	♫ SH: *There isn't anytime to waste then.* **p, S, LP** ⟷	SH: A więc nie ma czasu do stracenia. [BT: So there is no time to lose.]	SH: A więc nie ma czasu do stracenia, [BT: So there is no time to lose.,
4	1:38:41		CP: stationary (horizontal) CD: CS, SA VC: LK, pipe, smoke ED: establishing shot VF: Sherlock observes Coward.	◎ △ ◇ ⊟ (><) ⟹ Sherlock sits calmly on a chair smoking his pipe. He tilts his head to right side, keeps his eyes wide open expressing his 'bank face'.	♫ SH: *Is there?* **p, S, LP, #** ⟷	SH: X	SH: prawda? [BT: Is there?]

SH2_1

Frame No	Time (hh:mm:ss)	Frame	Visual Image	Kinesic Action	Soundtrack	Voice-over	Subtitles
1	0:12:41		CP: stationary (vertical) CD: MLS, SA VC: props, settings, costumes ED: establishing shot VF: Sherlock does not keep eye contact with Watson and looks through the window.	Sherlock stands straight looking through the window.	♫ SH: Would you prefer I joined you in the fashion faux pas... n, F, LP ↔	x	x
2	0:12:42		CP: stationary (vertical) CD: MLS, SA VC: props, settings, costumes ED: establishing shot VF: Sherlock looks at Watson reading a newspaper.	Sherlock stands straight and turns his sight, head and body towards Watson.	♫ SH: Would you prefer I joined you in the fashion faux pas... n, F, LP ↔	SH: Wolisz taki nietakt.. [BT: Would you prefer such a gaffe...]	SH: Wolałbyś taki nietakt... [BT: Would you prefer such a gaffe...]
3	0:12:43		CP: stationary (vertical) CD: MLS, SA VC: props, settings, costumes ED: establishing shot VF: Sherlock looks at Watson reading a newspaper.	Sherlock shrugs his shoulders and points with his head and cloth on Watson.	♫ SH: Would you prefer I joined you in the fashion faux pas... n, F, LP ↔	SH: Wolisz taki nietakt... [BT: Would you prefer such a gaffe...]	SH: Wolałbyś taki nietakt... [BT: Would you prefer such a gaffe...]
4	0:12:44		CP: stationary (horizontal) CD: MCS, LA VC: props, settings, costumes ED: establishing shot VF: Sherlock keeps eye contact with Watson.	Sherlock turns his sight, head and body towards Watson. Sherlock keeps his head straight.	♫ SH: Would you prefer I joined you in the fashion faux pas... n, F, LP ↔	SH: Wolisz takie nietakt, [BT: Do you prefer such a gaffe...]	SH: Wolałbyś takie nietakt, [BT: Would you prefer such a gaffe...]
5	0:12:45		CP: stationary (horizontal) CD: MCS, LA VC: props, settings, costumes ED: establishing shot VF: Sherlock looks at Watson's appearance.	Sherlock lowers his head and looks at Watson from above.	♫ SH: ...of wearing fine military dress with that heinous handmade scarf... n, F, LP ↔	SH: ...jak schludny strój wojskowego w połączeniu ze szpetnym szalikiem... [BT: ...a neat military outfit in combination with an ugly scarf...]	SH: ...jak porządny strój wojskowego ze szpetnym szalikiem, [BT: like a decent outfit of a military man with hideous scarf...]
6	0:12:46		CP: stationary (horizontal) CD: MCS, LA VC: props, settings, costumes ED: establishing shot VF: Sherlock looks at Watson's appearance.	Sherlock lowers his head and looks at Watson from above. Sherlock distorts his mouth with a grimace.	♫ SH: ...of wearing fine military dress with that heinous handmade scarf... n, F, LP ↔	SH: ...jak schludny strój wojskowego w połączeniu ze szpetnym szalikiem... [BT: ...a neat military outfit in combination with an ugly scarf...]	SH: jak porządny strój wojskowego ze szpetnym szalikiem, [BT: ...like a decent outfit of a military man with hideous scarf...]

|						
7	0:12:47		CP: stationary (horizontal) CD: MCS, LA VC: props, settings, costumes ED: establishing shot VF: Sherlock closes his eyes.	Sherlock keeps head lowered and closes his eyes. Sherlock expresses his 'blank face'.	♫ SH: ...of wearing fine military dress with that heinous handmade scarf... **n, F, LP** ↔	SH: ...w połączeniu ze szpetnym szalikiem [BT: ...a neat military outfit in combination with an ugly scarf...]	SH: ...jak porządny strój wojskowego ze szpetnym szalikiem, [BT: like a decent outfit of a military man with hideous scarf...]
8	0:12:48		CP: stationary (horizontal) CD: MCS, LA VC: props, settings, costumes ED: establishing shot VF: Sherlock divers his sight from Watson and keeps looking through the window.	Sherlock raises his head and direct his sight from Watson again. Sherlock keeps his eye wide opened expressing his 'blank face'.	♫ SH: ...clearly one of your fiancée's early efforts? **n, F, LP, #** ↖	SH:... wydzierganym przez narzeczoną? [BT: knitted by your fiancee?]	SH: ...dziełem rąk twojej narzeczonej? [BT: ... handmade by your fiancee?]
9	0:12:50		CP: stationary (horizontal) CD: MCS, LA VC: props, settings, costumes ED: establishing shot VF: Sherlock keeps looking through the window.	Sherlock raises his head and direct his sight from Watson again. Sherlock keeps his eye wide opened eyebrows expressing his 'blank face'.	♫ SH: ...clearly one of your fiancée's early efforts? **n, F, LP, #** ↖	SH: ... wydzierganym przez narzeczoną? [BT: ...knitted by your fiancee?]	SH: ...dziełem rąk twojej narzeczonej? [BT: ... handmade by your fiancee?]
10	0:12:52		CP: stationary (horizontal) CD: MCS, LA VC: props, settings, costumes ED: establishing shot VF: Sherlock directs his eyes again to Watson.	Sherlock lowers his head again and looks at Watson from above. Sherlock smirks while talking.	♫ SH: ...clearly one of your fiancée's early efforts? **n, F, LP, #** ↖	SH: ... wydzierganym przez narzeczoną? [BT: ...knitted by your fiancee?]	SH: ...dziełem rąk twojej narzeczonej? [BT: ... handmade by your fiancee?]

Frame No	Time (hh:mm:ss)	Frame	Visual Image	Kinesic Action	Soundtrack	Voice-over	Subtitles
1	0:02:01		CP: stationary (horizontal) CD: MCS, SA VC: settings, props, costumes ED: continuous VF: Mycroft is looking at Watson from above.	◇△☺ (><) ⟷ Mycroft is looking at Watson, frowning and raising his eyebrows. He keeps his head straight. Mycroft wants to drink Watson's health.	♫ MC: Your very good health, doctor. **n, F, MP** ↔	MC: Pańskie zdrowie, doktorze. [BT: Your health, doctor.]	MC: Pańskie zdrowie, doktorze. [BT: Your health, doctor.]
2	0:22:03		CP: stationary CD: CS SA VC: - ED: continuous VF: Sherlock keeps a direct eye contact with Watson	△◻⊖ (><) ⟷ Mycroft raises his head up looking at Watson from above. Mycroft frowns and raises his eyebrows.	♫ MC: Your very good health, doctor. **n, F, MP** ↔	MC: Pańskie zdrowie, doktorze. [BT: Your health, doctor.]	MC: Pańskie zdrowie, doktorze. [BT: Your health, doctor.]
3	0:22:04		CP: stationary (horizontal) CD: MCS, SA VC: settings, props, costumes ED: continuous VF: Mycroft is looking at Watson from above.	△ ⊖ ◻ ◇ (><) Mycroft lowers his head again looking at Watson from above. Mycroft frowns and raises his eyebrows.	♫ MC: Shame none of your friends could make it. **n, F, MP, #** ↖	MC: Szkoda, że przyjaciele nie przyszli. [BT: Shame your friends did not come.]	MC: Szkoda, że przyjaciele nie przyszli. [BT: Shame your friends did not come.]
4	0:22:05		CP: stationary (horizontal) CD: MCS, SA VC: settings, props, costumes ED: continuous VF: Mycroft is looking at Watson from above.	◇ △ (><) ⊖ Mycroft turns his head a bit left looking at Watson from above. Mycroft frowns and raises his eyebrows.	♫ MC: Shame none of your friends could make it. **n, F, MP, #** ↖	MC: Szkoda, że przyjaciele nie przyszli. [BT: Shame your friends did not come.]	MC: Szkoda, że przyjaciele nie przyszli. [BT: Shame your friends did not come.]
5	0:22:06		CP: stationary (horizontal) CD: MCS, SA VC: settings, props, costumes ED: continuous VF: Mycroft is looking at Watson from above.	△ ⊖ (><) ◻ Mycroft lowers and his head again on the right side looking at Watson from above. Mycroft frowns and raises his eyebrows.	♫ MC: Shame none of your friends could make it. **n, F, MP, #** ↖	MC: Szkoda, że przyjaciele nie przyszli. [BT: Shame your friends did not come.]	MC: Szkoda, że przyjaciele nie przyszli. [BT: Shame your friends did not come.]
6	0:22:07		CP: stationary (horizontal) CD: CS, SA VC: settings, props, costumes ED: continuous VF: Sherlock keeps eye contact with Mycroft.	△ ◻ ⊗ ⊖ ◻ Sherlock starts laughing when heard Mycroft speaking to Watson. Sherlock turns his head towards Mycroft and raises his eyebrows.	♫ ☼	X	X

7	0:22:08		CP: stationary (horizontal) CD: CS, HA VC: settings, props, costumes ED: continuous VF: Watson looks at Mycroft and Sherlock.	⬒ ⬭ Watson is sitting on the chair keeps looking at Sherlock and Mycroft and smirks.	♫ ✿		X	X
8	0:22:09		CP: stationary (horizontal) CD: CS, SA VC: settings, props, costumes ED: continuous VF: Sherlock directs his sight towards Watson.	⬒ ⊗ ⬭ Sherlock starts laughing and smirking turing his head towards Watson.	♫ ✿		X	X
9	0:22:10		CP: stationary (horizontal) CD: MCS, SA VC: settings, props, costumes ED: continuous VF: Mycroft looks at Watson.	⬒ ⊗ ⟷ Mycroft keeps smirking and laughing at Watson. He keeps a glass in his right hand.	♫ ✿		X	X
10	0:22:11		CP: stationary (horizontal) CD: MCS, SA VC: settings, props, costumes ED: continuous VF: Mycroft looks at Watson.	⊗ ⬭ ⬒ ⟷ Mycroft pulls his head down again looking at Watson from below and smirks at him. Mycroft raises his hand with a glass to drink Watson's health.	♫ ✿		X	X
11	0:22:12		CP: stationary (horizontal) CD: CS, HA VC: settings, props, costumes ED: continuous VF: Watson is looking at Mycroft and Sherlock.	⬒ ⊗ ⬭ (><) △ Watson raises his head, and raises his eyebrows significantly. Watson is laughing loudly with a wry smile.	♫ ✿		X	X

243

SH2_3

Frame No	Time (hh:mm:ss)	Frame	Visual Image	Kinesic Action	Soundtrack	Voice-over	Subtitlig
1	0:45:50		CP: stationary (vertical, crane shot) CD: CS, SA VC: action (shooting), settings, costumes, props like pipe ED: continuous VF: Sherlock is looking at Watson shooting to soldiers.	◇ ⟷ (><) ▢ Sherlock is lying on the floor with his hands behind the head. In the relaxed pose he is looking at Watson frowning and showing his 'blank face'.	♫ ☼	X	X
2	0:45:57		CP: stationary (vertical, crane shot) CD: CS, SA VC: action (shooting) settings, costumes, props ED: continuous VF: Sherlock is looking at Watson while they are lying	✕ ⊟ ⟷ Sherlock is turned directly to Watson on the right side. He opens is mouth widely and shout at Watson. Sherlock is trying to look at Watson from above despite the chips flying	♫ ☼ SH: *I said make it count. How many windows must I provide?* **ff, F, LP, #**	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]
3	0:45:58		CP: stationary (vertical, crane shot) CD: CS, SA VC: action (shooting), settings, costumes, props ED: continuous VF: Sherlock is looking at Watson while they are lying on the floor.	✕ ⊟ ⟷ Sherlock moves his face closer to Watson's face. Sherlock keeps shouting and looking at Watson from above while Watson still covers his face with hands.	♫ ☼ SH: *I said make it count. How many windows must I provide?* **ff, F, LP, #** ⟷	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]
4	0:45:59		CP: stationary (vertical, crane shot) CD: CS, SA VC: action (shooting), settings, costumes, props ED: continuous VF: Sherlock is looking at Watson while they are lying on the floor.	✕ ⊟ ⟷ Sherlock moves his face even closer to Watson's face. Sherlock keeps shouting and looking at Watson from above while Watson still covers his face	♫ ☼ SH: *I said make it count. How many windows must I provide?* **ff, F, LP, #** ⟷	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]
5	0:46:00		CP: stationary (vertical, crane shot) CD: CS, SA VC: action (shooting), settings, costumes, props like pipe ED: continuous VF: Sherlock is looking at Watson while they are lying on the floor.	◇ ⊟ ⟷ Sherlock moves his head up from the floor to look at Watson from a distance. Sherlock keeps his mouth tightly closed showing his 'blank face'. Watson still covers his face with hands.	♫ ☼ SH: *I said make it count. How many windows must I provide?* **ff, F, LP, #** ⟷	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]
6	0:46:01		CP: stationary (vertical, crane shot) CD: CS, SA VC: action (shooting), settings, costumes, props like pipe ED: continuous VF: Sherlock is trying to keep eye contact with Watson.	⊟ ⟷ (><) ✕ Sherlock lies his head again on the floor patronising Watson. Sherlock starts shouting again. Watson still covers his face with hands.	♫ ☼ SH: *I said make it count. How many windows must I provide?* **ff, F, LP, #** ⟷	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]
7	0:46:02		CP: stationary (vertical, crane shot) CD: CS, SA VC: action (shooting), settings, costumes, props like pipe ED: continuous VF: Sherlock is closing his eyes.	⊟ ⟷ Sherlock closes his eyes and calms down, Watson moves his hand from his face.	♫ ☼ SH: *I said make it count. How many windows must I provide?* **ff, F, LP, #** ⟷	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]	SH: Mówiłem celnie! Ile szans mam ci dawać? [BT: I said accurately! How many chances do I have to give?]

Appendix 6 - Eye-tracking Data Quality

Subject	Plec	Tracking Ratio [%]	Subject	Plec	Tracking Ratio [%]
P01A	k	78.4	P15A	k	81.2
P02A	k	72.1	P17B	k	86.8
P01B	k	78.6	P18B	m	78.1
P02B	k	71.2	P16A	m	78.7
P03A	k	73.0	P17A	k	84.4
P04A	k	81.1	P19B	k	91.7
P03B	k	78.8	P18A	k	80.1
P05A	k	71.0	P19A	k	89.8
P05B	k	74.9	P20A	m	92.3
P06B	m	82.1	P20B	k	73.1
P06A	k	70.4	P21B	k	76.0
P07A	k	71.3	P22B	k	93.7
P07B	m	77.4	P21A	m	72.4
P08B	k	71.1	P24B	m	73.4
P08A	k	77.7	P22A	k	73.9
P10B	k	79.1	P25B	k	74.9
P09A	k	90.4	P26A	k	81.5
P10A	k	78.3	P27B	k	71.8
P11B	m	78.7	P30A	k	75.7
P13A	k	79.5	P32B	m	78.6
P14A	m	71.4	P33B	k	71.7

Appendix 7 - Detailed Eye-Tracking Data

(please see memory stick for information)

Appendix 8 - Questionnaire Responses

(please see memory stick for information)