

Relevant abuse?

**Investigating the effects of an abusive subtitling
procedure on the perception of TV anime using
eye tracker and questionnaire**

Colm Caffrey

Thesis submitted for the Degree Doctor of Philosophy

School of Applied Language and Intercultural Studies
Dublin City University

Supervisor: Dr. Minako O'Hagan

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of doctor of philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____

ID No.: 50708500

Date: DATE

Table of contents	
Abstract.....	x
List of publications and presentations	xi
Chapters in books and refereed articles	xi
Conference papers	xi
Invited talk.....	xi
Translation	xi
Acknowledgements	xii
CHAPTER 1.INTRODUCTION	1
1.1 Research context	2
1.2 Scope of the thesis	5
1.3 Thesis structure	6
CHAPTER 2.ABUSIVE SUBTITLES.....	9
2.1 Definition of abusive subtitles.....	10
2.2 Constraints in subtitling	12
2.3 The translator's (in)visibility	15
2.4 Abusive subtitles research and fansubs.....	17
2.4.1 Categorisation of abusive procedures in fansubs	18
2.5 An abusive procedure: The AD Vid-note.....	19
2.6 Translating the visual in AVT	21
2.7 Abusive subtitles on commercial TV anime DVD	23
2.7.1 Technological factors	23
2.7.2 Distributor-centred factors.....	24
2.7.3 Fan-centred factors.....	25
2.8 Summary.....	27
CHAPTER 3.TV ANIME ON DVD	28
3.1 TV anime: A definition	29
3.2 Anime industry and production processes	30
3.3 Anime genres.....	33
3.4 Anime and cultural knowledge	34
3.5 The translation of TV anime	36
3.5.1 Transfer of character voice in dubbed and subtitled anime.....	37
3.5.2 Transfer of word play in anime	37
3.5.3 Dealing with pictorial elements.....	38
3.6 TV anime as a source text	43

3.7 Perception of TV anime on DVD	44
3.7.1 Culturally marked visual nonverbal cues: Introduction	45
3.8 Summary.....	48
CHAPTER 4.A RELEVANCE THEORETICAL APPROACH TO THE PERCEPTION OF SUBTITLED AV CONTENT	49
4.1 Relevance theory	50
4.1.1 Positive cognitive effects.....	51
4.1.2 Processing effort.....	53
4.2 Summary.....	57
CHAPTER 5.MEASURING PROCESSING EFFORT USING AN EYE TRACKER.....	58
5.1 Using eye tracking data to measure processing effort	59
5.1.1 Fixation-based measurements.....	61
5.1.2 Pupillometric measurements.....	64
5.2 Eye tracking tool: <i>Tobii 1750</i> eye tracker	71
5.3 Eye tracking software: <i>ClearView</i>	72
5.3.1 Recording data using ClearView.....	72
5.3.2 Data output	73
5.4 Summary.....	77
CHAPTER 6.EMPIRICAL PERCEPTION STUDIES IN AVT	78
6.1 Perception studies in AVT: Review	81
6.2 Studies employing consecutive data collection	82
6.2.1 Questionnaire.....	82
6.2.2 The questionnaire as a form of data collection.....	87
6.3 Studies employing only simultaneous data collection	89
6.3.1 Eye tracker only	89
6.3.2 Using eye tracker only as a method of data collection	92
6.4 Studies employing simultaneous and consecutive data collection	93
6.4.1 Interactive task and questionnaire	93
6.4.2 The “protest button” as a form of data collection.....	94
6.4.3 Observation and questionnaire	95
6.4.4 Observation as a form of data collection.....	95
6.4.5 Eye tracking and questionnaire.....	97
6.4.6 Eye tracker and questionnaire as a method of data collection.....	99
6.5 Sample size in reviewed AVT perception studies	100

6.6 Summary.....	101
CHAPTER 7.METHODOLOGICAL CONSIDERATIONS AND EXPERIMENT DESIGN.....	103
7.1 Pilot study	104
7.1.1 Aim.....	104
7.1.2 Subjects	105
7.1.3 Material	105
7.1.4 Questionnaire design	106
7.1.5 Procedure	106
7.1.6 Issues arising from pilot study.....	107
7.2 Hypothesis, independent and dependent variables	109
7.2.1 Questionnaire data: Independent variables	109
7.2.2 Questionnaire data: Dependent variables.....	110
7.2.3 Fixation-based and pupillometric data: Independent variables	111
7.2.4 Fixation-based and pupillometric data: Dependent variables.....	112
7.3 Units of analysis.....	114
7.4 Statistical testing	114
7.4.1 Statistical test used with questionnaire data: The paired sample t-test	116
7.4.2 Statistical test used with fixation-based and pupillometric data: Two factor within-subject analysis of variance (ANOVA).....	118
7.5 Experiment validity	119
7.5.1 Measurement validity	119
7.5.2 Internal validity	120
7.5.3 External validity.....	127
7.6 Experiment design	128
7.6.1 Source text selection.....	129
7.6.2 Ethical approval	131
7.6.3 Method.....	131
7.6.4 Procedure	139
7.7 Summary.....	140
CHAPTER 8.DATA ANALYSIS	141
8.1 Questionnaire data.....	142
8.1.1 Hypotheses about questionnaire data.....	142
8.1.2 PCEs relating to CVNCs	144

8.1.3 Declared PCEs regarding CVNCs	148
8.1.4 PCEs relating to subtitle.....	151
8.1.5 Processing effort.....	154
8.1.6 Summary of questionnaire analysis	156
8.2 Fixation-based and pupillometric data.....	157
8.2.1 Hypotheses about fixation-based and pupillometric data.....	158
8.2.2 Percentage of skipped subtitles	160
8.2.3 Percentage gaze time in subtitles	161
8.2.4 Mean fixation duration in subtitles.....	162
8.2.5 Word fixation probability.....	165
8.2.6 Median pupil size	166
8.2.7 Summary of fixation-based and pupillometric results.....	167
8.3 Summary.....	168
CHAPTER 9.CONCLUSION	172
9.1 Research questions revisited	173
9.1.1 Question 1	175
9.1.2 Question 2	177
9.2 Strengths and limitations of the present study	178
9.3 Directions for future work	181
References.....	184

Appendix A: Preparation of stimuli for use with *ClearView*

Appendix B: Pilot study questionnaire

Appendix C: Mean values for fixation-based and pupillometric dependent variables

Appendix D: Email from *Tobii* about screen resolution

Appendix E: Content questionnaire from experiment

Appendix F: Recruitment email

Appendix G: Email questionnaire

List of figures

Figure 1.1. Still illustrating abusive procedures from a commercially produced TV anime DVD	3
Figure 2.1. Screen capture from Episode 24 of the VH1 programme <i>Pop-up Video</i>	20
Figure 2.2. Extras menu on <i>Paniponi dash!</i> where pop-up gloss can be selected	21
Figure 3.1. Example of the production cycle for anime	30
Figure 3.2. Example of mimetic expression in TV anime	41
Figure 3.3. Two different procedures for translating mimetic expressions	42
Figure 3.4. Screen capture illustrating part of the semiotic mix present in TV anime	45
Figure 4.1. Nosebleed indicating arousal	52
Figure 5.1. Waveform showing maximum pupil diameter calculated using trial-aggregated fine-grained method.....	68
Figure 5.2. <i>Tobii 1750</i> eye tracker	71
Figure 5.3. Example of "good" and "bad" pupil calibration in <i>ClearView</i>	73
Figure 5.4. Screen capture from <i>ClearView</i> statistics template in <i>Excel</i>	74
Figure 5.5. Example of hotspot image	75
Figure 6.1. Screen capture showing a "dual speaker" subtitle	90
Figure 7.1. Screen capture from the same scene in <i>treatment</i> condition and <i>normal</i> condition	110
Figure 7.2. Screen captures depicting an example from each experiment condition described in Table 7.1	112
Figure 7.3. Screen capture from the same point of an excerpt with full captions and normal subtitles.....	122
Figure 7.4. Percentage gaze time in pop-up gloss area in order of appearance of pop-up gloss for selected subjects.....	126
Figure 7.5. Examples referred to in Table 7.4 of pop-up gloss explaining culturally marked items in each semiotic channel	130
Figure 7.6. Screen captures of CVNCs from each excerpt	133
Figure 8.1. Number of correct answers to CVNC question for each excerpt according to experiment condition	145
Figure 8.2. Screen captures from <i>Paniponi Dash!</i> and <i>The Living Daylights</i>	146

Figure 8.3. Declared understanding of CVNC in each excerpt according to experiment condition and difference between conditions.....	149
Figure 8.4. Actual and declared PCEs related to CVNCs for each excerpt according to experiment condition	150
Figure 8.5. Number of correct answers to subtitle question for each excerpt according to experiment condition and difference between conditions .	151
Figure 8.6. Screen capture of subtitles from Excerpt 5 and Excerpt 7 on which subtitle content question was based	152
Figure 8.7. Screen captures from Excerpt 6 showing rabbit and cat characters	154
Figure 8.8. Number of subjects who declared subtitle speed "fast" for each excerpt according to experiment condition.....	155
Figure 8.9. Mean percentage of skipped subtitles according to experiment condition	161
Figure 8.10. Mean percentage gaze time in subtitle area according to experiment condition.....	162
Figure 8.11. Mean fixation duration (ms) in subtitle area according to experiment condition.....	163
Figure 8.12. Mean word fixation probability according to experiment condition	165
Figure 8.13. Mean median pupil size according to experiment condition.....	167
Figure 8.14. Fixation heatmap for same scene in normal and treatment conditions.....	171

List of tables

Table 5.1. Mean pupillary response calculated using time-aggregated coarse method.....	69
Table 5.2. Extract from <i>Excel</i> spreadsheet showing fixation count for selected subjects.....	75
Table 6.1. Same subtitle in high frequency and low frequency conditions	97
Table 6.2. Same subtitle in low-cohesion and high-cohesion conditions	98
Table 7.1. Experiment conditions used with fixation-based and pupillometric data.....	111
Table 7.2. Results of Shapiro Wilk W tests for normal distribution of questionnaire data	118
Table 7.3. Percentage eye data present for each subject and excerpt	124
Table 7.4. Pop-up gloss in each of the semiotic channels	130
Table 7.5. Screen capture and pop-up gloss description of CVNC on which CVNC question was based for each excerpt	134
Table 7.6. Experiment condition in which subjects saw each excerpt.....	140
Table 8.1. Subtitles from Excerpt 1	146
Table 8.2. Subtitle from Excerpt 11	147
Table 8.3. Subtitles from Excerpt 5	147
Table 8.4. Number of subtitles in each condition for subjects 1-10 and 11-20 respectively.....	157
Table 8.5. Total number of data entries used in fixation-based and pupillometric analyses according to experiment condition	157
Table 8.6. Approximate mean fixation duration in several tasks	164

Abstract

The storage capacity of DVD means multiple subtitle streams can be included on one disc. This has allowed some producers to include subtitle streams with experimental procedures that we will term as “abusive” subtitles (Nornes 1999). Abusive subtitles break subtitling norms in an attempt to be more faithful to the source text and increase the translator’s visibility. This thesis focuses on one such abusive procedure, namely the pop-up gloss. It refers to pop-up notes that explain culturally marked items appearing in each of the semiotic channels. Already popular with amateur subtitlers of anime (Japanese animation), pop-up gloss has come to percolate into commercially released anime DVD.

This thesis investigates the question as to what effect the use of pop-up gloss has on viewer perception of TV anime in terms of positive cognitive effects (PCEs) and processing effort. A second question seeks to ask the validity of pupillometric measurements for measuring the processing effort experienced while viewing subtitled AV content. A novel methodology is applied where PCEs are measured using traditional questionnaire data, while processing effort is measured using a combination of questionnaire-based data, and fixation-based and pupillometric data gathered with an eye tracker.

A study with 20 subjects indicates that the use of pop-up gloss does increase the PCEs experienced by subjects regarding items the pop-up gloss describes, while more processing effort is required by viewers when pop-up gloss is used. The analysis of pupillometric data suggests that they are suitable for measuring processing effort during the viewing of subtitled AV content.

Key words: Perception, eye tracking, audiovisual translation, anime

List of publications and presentations

The following lists publications, conference papers, an invited talk and translation that originated in the research carried out for this thesis.

Chapters in books and refereed articles

Caffrey, C. 2008a. Viewer perception of visual nonverbal cues in subtitled TV Anime. *European Journal of English Studies*. 12(2), pp163 – 178.

Caffrey, C. 2008c. Using pupillometric, fixation-based and subjective measures to measure the processing effort experienced when viewing subtitled TV anime with pop-up gloss. *IN: Göpferich, S., Jakobsen, A.L. and Mees, I.M. (eds.) Looking at Eyes. Eye-Tracking Studies of Reading and Translation Processing. Copenhagen Studies in Language*, 36, pp125 – 144.

Caffrey, C. forthcoming. Eye tracking application for measuring the effects of experimental subtitling procedures on viewer perception of subtitled AV content. *IN: d'Ydewalle, G. and Perego, E. (eds.) As yet untitled volume on subtitling and eye tracking.*

Conference papers

Caffrey, C. 2007. *Innovative Insanity? “Abusive” Subtitling Strategies and their Effects on Viewers’ Reading Times*. Intermedia Conference, Łódź, Poland, 13-14 April 2007.

Caffrey, C. 2008b. *Relevant Abuse? Effects of Abusive Subtitling Strategies on Viewer Perception of TV Anime*. Audiovisual Translation: Multidisciplinary Approaches Conference, Montpellier, France, 19-20 June 2008.

Invited talk

Caffrey, C. 2008d. *Processing subtitled AV content*. Keynote paper at I Seminar on Eye-Tracking and Audiovisual Translation. International seminar organised by Transmedia Catalonia Research Group and the Universitat Autònoma de Barcelona, Barcelona, Spain, 17 – 18 November 2008.

Translation

Ōishi, K. (2005) 死人を恋う. Tokyo: Kobunsha. *Translated title “In love with the dead”*. Translated from the Japanese to English for script development purposes for SPFilms.

Acknowledgements

This thesis would never have been possible without the many people who supported me through the process. Among those, there are a few I would like to thank especially. First and foremost, I would like to thank my supervisor, Minako O'Hagan. Minako was a fantastic supervisor, and provided me with constant support, feedback and constructive comments throughout the whole thesis process. I cannot say how grateful I am for the effort she put into my supervision from start to finish, and for the friendly atmosphere she always created in our dealings.

Dorothy Kenny's valuable feedback from her revision of a draft of my thesis is greatly appreciated. Also, thanks to everyone in the postgrad room for their support and friendship. Thanks too to the twenty anonymous subjects who kindly gave their time, effort and eye recordings for the research.

Thanks to Sharon O'Brien and Delia Chiaro for taking time out of their busy schedules to be the examiners for the thesis, and to Sharon for securing the eye tracker for SALIS, without which it would not have been possible for me to carry out the research. I would also like to thank the Irish Research Council for the Humanities and Social Sciences for their funding support, which enabled me to focus my efforts fully on my research.

Finally I would like to thank my parents, whose support throughout my life has led me to all the opportunities that I have had. They have always been there for me, and knowing that has given me the sense of security to do everything I have. I would also like to thank Mick Mullen for keeping my feet on the ground during the process, and for being there to celebrate and commiserate. Thanks for the support and always being there to listen and encourage me. I would like to thank Elisa Alonso Sanz for her help with the statistical analyses carried out in the thesis. My thanks also go to Leire Zirrión Corres for checking my translations of Spanish-language quotes, and her encouragement over the last few difficult months of the thesis write-up.

And sorry to everyone for any and all rants – no more...hopefully.

THESIS CONVENTIONS

- IN INSTANCES WHERE THE AMERICAN ENGLISH SPELLING OF A WORD IS USED IN QUOTATIONS, IT HAS BEEN CHANGED TO THE BRITISH ENGLISH SPELLING TO MAINTAIN CONSISTENCY IN THE THESIS.
- WHEN AN ANIME OR MANGA IS NAMED FOR THE FIRST TIME, IT IS PRESENTED IN THE FOLLOWING FORMAT “JAPANESE TITLE (*ROMANIZATION OF ORIGINAL TITLE*, ENGLISH TRANSLATION). IN SUBSEQUENT INSTANCES, THE ENGLISH TRANSLATION OF THE TITLE WILL BE USED.
- THE TERM POSITIVE COGNITIVE EFFECT IS ABBREVIATED FOR THE MAJORITY OF THE THESIS AS PCE. HOWEVER, IT IS USED IN FULL FOR THE ENTIRETY OF CHAPTER 4.
- WHEN [AT] IS SEEN IN THE THESIS, IT INDICATES THAT THE QUOTATION WAS TRANSLATED BY THE AUTHOR. THE TRANSLATION IS SHOWN IN ITALICS.

CHAPTER 1. **INTRODUCTION**

1.1 Research context

Over the past two decades, the media on which audiovisual (AV) content is viewed have undergone several changes. From the arrival of DVD technology to the home video market in the late 1990s, to the more recent emergence of the next generation DVD format, Blu-ray, the storage capacity available to distributors of AV content has increased drastically since the days of VCR.

While VCR presented distributors with the need to choose between a subtitled or dubbed version of foreign AV content for each video cassette, the enhanced storage capacity of DVD allowed distributors to fit multiple audio and subtitle tracks on one disc. Blu-ray, which appears to have won the battle for domination of the next generation DVD market over HD-DVD, offers even more storage capacity. However, none of the above formats offer as much capacity for storage and accessibility as the internet, which as a form of distributed storage, enables anyone with a broadband internet connection to watch, download and upload a more wide and varied range of AV content than was ever previously possible.

The advent of Web 2.0, a collection of concepts relating to general developments in web interaction, has had an influence on the way that web users create and share digital content online. Some of the key themes of Web 2.0, such as open access and distribution, user-generated content, and decentralisation (Dean 2007) have had a particularly marked effect on the creation, translation and distribution of AV content online. The ease with which untrained and amateur subtitlers can subtitle, edit and upload AV content has brought with it an increase in the exposure of internet users to what are referred to in a positive light as “abusive subtitles” (Nornes 1999, 2004). Abusive subtitles are subtitles that experiment and contravene current subtitling norms to provide a more “authentic” translation, foregrounding the work as a translation and at the same time, the translator’s work.

While abusive subtitles can be found in a variety of text-types, there is no doubt that anime, or Japanese animation, is one of the most popular genres where abusive subtitles may be found. This can be linked back to the days of VHS

and anime fan communities, when fans would translate anime that were unlikely to be made available in an English language version. These fan produced subtitles, or fansubs, have since been an integral part of anime fan communities, and provide a means to create what fans consider more “authentic” translations than those available on commercially released anime (Cubbison 2005). Some types of abusive subtitles used in both fansubs and commercially released anime to increase the authenticity of a translation include the use of different fonts to match the mood of a piece (e.g. gothic font for a period vampire drama), subtitles exceeding two lines in length, and pop-up glosses to explain culturally marked content (see Figure 1.1). The latter procedure, pop-up gloss, could theoretically be used to aid non-Japanese viewers’ perception of anime texts, where “a product encoded by and for one culture is decoded by a very different culture” (Levi 2006 p44).



Figure 1.1. Still illustrating abusive procedures from a commercially produced TV anime DVD (*Paniponi dash!* 2007)

While it is clear that the use of pop-up gloss is likely to have an effect on perception of AV content, the characteristics or extent of this effect have yet to be empirically proven. Nevertheless, the adoption of the pop-up gloss by some commercial anime DVD producers such as *ADV Films* indicates that certain producers entertain the notion that the use of pop-up gloss will have a positive effect on viewers’ perception of AV content in certain contexts.

The present research forms part of a current trend in audiovisual translation (AVT) research that aims to go some way towards redressing the paucity of empirical studies on the perception of translated AV content highlighted by Antonini (2005 p210):

Considering the huge number of recipients of audiovisual translation products, why has audience perception of both dubbing and subtitling been so patently ignored?

Within this research movement, there has been a recent surge in interest in the use of data gathered using eye trackers¹. This interest is exemplified by the recent international seminar organised by the Transmedia Catalonia Research Group and the Universitat Autònoma de Barcelona in November 2008 and an upcoming volume to be edited by Elisa Perego and Géry d'Ydewalle, both with a focus on AVT and eye tracking. There has also been a recent increase in output from researchers into the perception of translated AV content employing data generated with eye trackers (Caffrey 2007, 2008a, 2008b, 2008c, Moran 2008, Perego and Del Missier 2008). The increased interest in the eye tracker as a research tool in AVT, and in empirical studies on the perception of translated AV content in general, adds to the motivation and the applicability of the present study to future AVT research.

While previous studies of the perception of translated AV content have largely dealt with the perception of humour (Fuentes Luque 2003, Chiaro 2004, Antonini *ibid*), the use of eye tracker as a data collection tool opens other avenues of research. While it is generally easy for a person to tell whether or not they found a joke funny, some internal processes can be more difficult to judge subjectively. One such process is the amount of effort it took to read a certain subtitle set, or how much time was spent in a certain part of the screen. It will be argued that the complementary use of data from the eye tracker with data from other sources allows the researcher to infer the properties of some internal processes with more confidence than the use of other more traditional data collection methods alone.

¹ An eye tracker is a device that records the movements of the eye, and in some cases, the size of the pupil.

1.2 Scope of the thesis

This thesis sets out to empirically investigate the effects that one abusive subtitling procedure, in the form of pop-up gloss applied in a TV anime on DVD, has on viewer perception of “culturally marked visual nonverbal cues” (CVNCs) and subtitles. It is an interdisciplinary study, and draws on factors from process-oriented research in translation studies (TS), AVT perception studies, intercultural studies as well as cognitive psychology. An experiment was developed drawing on the work of relevance theorists such as Sperber and Wilson (1986) and Yus (2008); AVT perception scholars such as Chiaro (2004) and Antonini (2005); and cognitive psychologists with experience using data collected with eye trackers such as d’Ydewalle and de Bruycker (2007) and Just and Carpenter (1993). The experiment sought to explore the following research questions:

1. What effect does the presence of pop-up gloss have on the perception of subtitled TV anime?
2. Can pupillometric data be used to accurately measure the processing effort experienced while viewing subtitled AV content?

The approach taken to perception in this study is based on Sperber and Wilson’s (ibid) relevance theory. Considering perception using their concepts of positive cognitive effects (PCEs) and processing effort, a methodology was developed to allow the application of these concepts in a systematic analysis of data gathered using a combination of methods. The eye movements and pupil diameter of 20 subjects were recorded as they viewed 11 excerpts from a popular TV anime series on DVD. The eye tracker data, as well as answers to questions posed after viewing each excerpt, provided the data for analysis.

The methodology employed in the present study was considered to allow for the empirical observation of the effects that the presence of pop-up gloss has on i.) the amount of accurate information retained by subjects, as well as ii.) the amount of processing effort they experience.

Finally, consideration is given to the suitability of one of the data collection methods used to measure processing effort in the thesis. Pupil size has been used as a method to measure processing effort in a number of prior studies,

ranging from reading sentences (Just and Carpenter 1993) to interpreting (Hyönä *et al.* 1995) and translating using translation memory software (O'Brien 2005, 2008). However, the present study is the first time that pupil size has been used to measure the processing effort experienced while viewing subtitled AV content. As such, the data provided by complementary methods of measuring processing effort is used to help determine the validity of the pupil size data as a means of measuring processing effort.

1.3 Thesis structure

The thesis consists of nine chapters. Following the current introduction chapter, Chapter 2 provides an overview of abusive subtitles, describing their attributes and the motivation behind their use. One of the translation constraints particularly relevant in the case of abusive subtitles, the visibility of the translator, is highlighted and the relevant research into abusive subtitling procedures is reviewed. The procedure being studied in this thesis, pop-up gloss, is explained in detail and the motivation behind some producers' use of abusive procedures on commercially released TV anime DVDs is explained.

Chapter 3 focuses on the source text genre used in the present study, TV anime on DVD. The chapter begins with a definition of the term TV anime in the context of the present study, followed by a brief background of the anime industry. The debate surrounding the use of translated anime as a tool for cultural education is briefly touched upon, and the relevant English-language research that has been carried out to date into the translation issues encountered in anime and manga, Japanese comics, is discussed. The chapter concludes with the decision to focus the present study on two aspects of the semiotic mix present in subtitled TV anime, culturally marked visual nonverbal cues (CVNCs) and subtitles.

Chapter 4 describes the relevance theoretical framework used to analyse the perception of subtitles and CVNCs, based on Sperber and Wilson's (1986, 1995) relevance theory. In particular, the two aspects that are used to analyse perception in the present study, positive cognitive effects (PCEs) and processing effort, are discussed. The tools used to measure PCEs and

processing effort in the current study, questionnaire and eye tracker, are then introduced.

Chapter 5 looks into the use of the eye tracker as a tool for measuring processing effort. It discusses the two types of eye data gathered in the present study, fixation-based and pupillometric data, in terms of their previous applications in studies using subtitled AV content, and their applicability in the present study. Four fixation-based measurements used in the present study are selected: percentage of subtitles skipped by the viewer, percentage gaze time in subtitle area, mean fixation duration in the subtitle area and word fixation probability. Furthermore, the decision for choosing a time-aggregated coarse method for dealing with pupillometric data rather than a trial-aggregated fine-grained method is discussed. The eye tracker that was used in the present study, the *Tobii 1750* eye tracker, is described. To conclude, some of the strengths and weaknesses of the software package used to deal with the eye tracker data, *ClearView*, are highlighted.

Chapter 6 provides a critical overview of the relevant prior empirical research on the perception of translated AV content. It also reviews the data collection methods used in prior research. The research review is divided into three categories, determined by the data collection method employed: consecutive data collection only, simultaneous data collection only, or simultaneous and consecutive data collection. The chapter concludes by highlighting two major factors that appear to have an impact on the sample size of empirical studies into the perception of translated AV content: scale of funding and data collection method.

Chapter 7 discusses the methodological considerations necessary to design the experiment. It describes the technological and experimental design issues arising from a pilot study carried out using questionnaire and eye tracker. The hypothesis, variables and statistical tests necessary to formulate a robust experimental design are then described. The validity of the experimental design is considered, paying particular attention to the challenges posed by the present study. The chapter concludes with a detailed description of the experiment carried out in the present study.

Chapter 8 reports the analyses of the data gathered from the experiment. The consecutively gathered data is discussed first, followed by the simultaneously gathered data. The results are categorised in terms of PCEs and declared PCEs relating to culturally marked visual nonverbal cues, PCEs relating to subtitles, and processing effort experienced by subjects while watching the excerpts.

Chapter 9 provides a summary of the analysis of the data and reflections by the researcher on the endeavour to empirically investigate the effects that pop-up gloss has on the perception of subtitled TV anime. It discusses the applicability of the present research methodology in future AVT perception studies, and the promising outlook for the applicability of pupillometry to study processing effort required when viewing subtitled AV content. The methodology is reviewed in terms of its strengths and weaknesses, and some possible avenues of future research are highlighted.

CHAPTER 2. **ABUSIVE SUBTITLES**

Rejoice - subtitles have been freed! For decades, they have been plain white lines of text tethered to the bottom of the screen in foreign films. But now...They're liberated (Beckman 2008).

The use of pop-up gloss in subtitled AV content is not a standard practice and can be considered as a form of “abusive” subtitling (Nornes 1999, 2004), i.e. it experiments with and differs from the current norms of subtitling. In this chapter, the notions of abusive translation and subtitling are discussed, and an overview is given of the relevant prior research that has been carried out to date into abusive subtitling. Possible reasons for abusive subtitling being used in commercially released DVD are discussed, with a particular focus on the TV anime market. A distinction is drawn between abusive strategies and procedures in translating AV content, and in conclusion, the motivation for selecting pop-up gloss as the abusive procedure for investigation in the present study is explained.

2.1 Definition of abusive subtitles

The term “abusive subtitling” was first used in an article by Nornes (1999), where he argues for the use of subtitle procedures that are experimental, both graphically and linguistically, as a method of challenging the status quo, which he refers to as being “corrupt”, in so far as the latter method of translation “conspires to hide its work – along with its ideological assumptions – from its own reader-spectators” (Nornes *ibid* p19). Nornes (*ibid*) goes on to encourage the use of “abusive” subtitles where the subtitler:

uses textual and graphic abuse – that is, experimentation with language and its grammatical, morphological, and visual qualities – to bring the fact of translation from its position of obscurity, to critique the imperial politics that ground corrupt practices while ultimately leading the viewer to the foreign original being reproduced in the darkness of the theatre.

Nornes (*ibid*) draws on Lewis (1985) who introduces the term “abusive translation”, in an article that is itself an abusive translation of an article by Lewis himself originally written in French (Lewis 1981). Lewis (1985) argues for the use of overt translation strategies that experiment with usage, introducing procedures atypical of the target language, rather than providing what he calls “servile” translations. The concept of abusive translation is based on the

premise that translation must transfer meanings from the original into a “framework that tends to impose a different set of discursive relations and a different construction of reality” (Lewis 1985 p36). So, when an utterance in Japanese is transferred into English, the message is made to suit the discursive structures of English. This may cause difficulties for the translator, especially if the translation must accurately interpret the original while remaining within the boundaries imposed by the target language and culture, and in the case of subtitling particularly, the medium.

Nornes (1999 p30) argues that current subtitling standards do not interpret the original source text accurately enough, as:

All that cannot be explained within the severe limits of the regulation subtitle gets excised or reduced to domestic meanings which are often irrelevant or inappropriate. These subtitlers claim to bring their readers/spectators to a pleasurable experience of the foreign...As for their audiences, they are kept ignorant of the conspiracy and the riches that remain hidden from the cinematic experience.

An example where the meaning of the text is manipulated to fit the target culture can be found in a scene from the commercial DVD release of the TV anime series *先生のお時間* (*sensei no ojikan*, Doki Doki School Hours 2005) as described in Caffrey (2008a p166):

A tall teenager hugs her short teacher. The teenager has previously explained her attraction to her teacher, partially because of the teacher’s petite stature, and as she hugs her, another student comments:

[Japanese dialogue]	あ 海老で鯛は釣れた
[English translation]	Ah, a bream was caught using a shrimp
[English subtitle]	Talk about a big fish and a little shrimp.

As the student says this, the image of a shrimp and a bream appear above the hugging characters. The Japanese dialogue is the equivalent of the English proverb ‘to use a sprat to catch a mackerel’. However, in this case, the image of the shrimp and bream would not allow for the direct translation of the dialogue in a subtitle. Instead, the subtitler explains the presence of the image by drawing attention to the size comparison and avoiding the proverbial meaning of the shorter character attracting the taller one.

In the case cited above the subtitler chose to alter the meaning of the source text rather than use a procedure that may have made his/her position as an

intermediary more overt. The most likely reason for this choice is the desire to ensure synchrony between the image and the subtitle. Although they do not occur frequently, in translation scenarios where the creator of the content is also the translator, the issues of synchrony and manipulation are less strict. This is because the creator/translator is able to freely “compensate” for any perceived loss in translation by providing additional commentary.

Appiah (2000) echoes Nornes’ (1999) call for more source-oriented explicit translation with what he calls “thick translation”, translation that includes annotations and glossaries to improve the understanding of culturally marked items in the source text, to be used in instances where the meaning can be difficult to convey to the target audience. There is some crossover between thick translation and abusive translation, as Hermans (2003 p387) illustrates when he argues that one of the advantages of thick translation is that “it advertises the fact that translation, interpretation and description are played out in the same discursive space”. Another similarity between thick translation and abusive translation can be seen in Hermans’ (ibid) description of thick translation:

as a highly visible form of translating it flaunts the translator's subject position, counteracting the illusion of transparency or neutral description, and instead introducing a narrative voice into the account, thereby arming it with an explicit viewpoint.

Both abusive translation and thick translation have the common feature that they foreground the translation as a translation, making the subtitler’s work more visible. These are forms of overt translation, or translation where “the function of the translation is to enable its readers to access the function of the original in its original linguacultural setting through another language” (House 1997 p29). However, as a form of constrained translation, subtitling does not traditionally aim to provide overt translations, where the work of the subtitler is clearly visible, but rather to conceal it.

2.2 Constraints in subtitling

Subtitling is considered as a form of constrained translation (Tifford 1982, Gottlieb 1992, Martí Ferriol 2007), where many of the difficulties encountered by the subtitler are imposed by the medium. A constrained translation is a form of

translation where the text is only one of the components of the message or is an intermediate stage in the production, as is the case in a song to be sung aloud. Mayoral *et al.* (1988) use Titford's (1982) concept of constrained translation and provide a description of several types of translation, noting the aspects of each text type that are sources of translational difficulty. Mayoral *et al.* (ibid p363) name two circumstances that will result in a translation being constrained:

- a. the existence of various systems of communication,
- b. in some cases, the change from the visual channel for the text in the SL to the aural channel for the reading or reciting of the message in the TL (adaptation of the message to the norms of aural language).

They outline five types of synchrony, defined as “the agreement between signals emitted for the purpose of communicating the same message” (Mayoral *et al.* ibid p359), and use these as a means of ranking different translation text types in terms of the constraints they incur.

Mayoral *et al.* (ibid) classify subtitles as the second most constrained form of translation after dubbing. They mention three aspects where the subtitler is constrained, content synchrony (which could possibly include music and image), spatial synchrony and synchrony of time. *Content synchrony* occurs when the meanings transmitted by different semiotic channels do not contradict each other or the whole message, *spatial synchrony* when semiotic channels occupy no more or less space than that which corresponds with them and *synchrony of time* is the agreement in time of different semiotic channels that communicate an item of information.

Developing the technical constraints listed by Mayoral *et al.* (ibid), Gottlieb (1992) categorises the constraints presented by subtitling under two headings: *textual*, or qualitative constraints and *formal*, or quantitative constraints. *Textual* constraints are generally speaking constraints imposed by the content, such as how a regional dialect should be indicated in translation, while *formal* constraints are imposed by the medium.

In the case of subtitling, the two primary formal constraints are time and space. Imposed on the subtitler by the AV medium, these constraints have led to an

“inescapable degree of prescriptivism” (Díaz Cintas 2004a p29), especially in cinema subtitling. Subtitlers are granted only a limited amount of time and space to display the subtitle on screen. The generally accepted norms for subtitles is that for a two-line subtitle, with 32 characters per line, a display time of six seconds should be used (d’Ydewalle and de Bruycker 2007). This falls in line with the six-second rule suggested by Baker (1982), which says that a two-line subtitle should be displayed for a maximum of six seconds and contain a maximum of 64 characters and spaces. Using this figure, display times for shorter subtitles can be calculated depending on the subtitle length. This method of timing is unusual from a cognitive studies perspective, as in most European languages words are not read letter by letter. According to eye-tracker based observations, some words, such as articles, are generally not fixated upon at all (d’Ydewalle and Gielen 1992).

Martí Ferriol (2007 p176) provides a classification of translation constraints present in AVT, which he categorises according to the phase of the AVT process in which they occur. In the translation phase, he lists five constraints: i.) *formal*, ii.) *linguistic*, iii.) *semiotic (or iconic)*, iv.) *socio-cultural* and v.) *“void”*.

- i.) The description of formal constraints essentially equals Gottlieb’s (1992), in that they concern the techniques and practices expected in subtitling and dubbing.
- ii.) Linguistic constraints are caused by the type of language used in the source language text, such as dialects, registers and spoken language.
- iii.) Semiotic (or iconic) constraints are incurred by signs transmitted in the nonverbal visual and aural channels (except in the case of songs).
- iv.) Socio-cultural constraints basically refer to the constraints caused by culturally marked items, and include verbal and visual cultural references.
- v.) The “void” constraint is defined as a “lack of presence (or absence) of constraints” (Martí Ferriol 2007 p176).

Before the translation phase, there is what Martí Ferriol (ibid) refers to as the preliminary phase. In this phase, there is only one constraint that he notes, the *professional* constraint. This constraint is imposed by the work conditions of the translator, and includes aspects such as the time restrictions, fees and style guides. One professional constraint typically imposed on the subtitler must be considered before abusive subtitling procedures may be used; that is the need to maintain the invisibility of the subtitler.

2.3 The translator's (in)visibility

One of the central issues in the debate concerning abusive subtitles is the culturally imposed constraint of the AV translator's invisibility. Of course, the idea of the translator's invisibility is by no means relegated only to the field of AVT. Referring to literary translation, Venuti (2006) remarks that we have been trained "in effect, to value translations with the utmost fluency, an easy readability that makes them appear untranslated...", while Díaz Cintas and Remael (2007 p139) cite the widely held notion that "the less attention [subtitles] draw to themselves, the less we notice them, and therefore the better they are". The resulting effect is that translations tend to go unnoticed unless there is a "bump in the surface" (Venuti 2006). These bumps are areas where the translator has made a decision that the receiver of the text considers to be unusual, and the work is recognisable as a translation, a phenomenon that Leppihalme (1997) refers to as "culture bumps". Like Lewis (1985) and Nornes (1999), Venuti (ibid) believes that a translation should be experienced differently than an original composition for the very reason that it is not an original, "because not only a foreign work, but a foreign culture is involved".

Díaz Cintas (2005 pp169-170) summarises some of the issues that revolve around the subtitler's visibility in an article where he describes the effects of DVD technology on subtitling:

La injerencia y presencia del traductor en el propio texto fílmico a través de intervenciones metatextuales, como notas a pie de página o glosas de determinados terminus, ha estado siempre fuera de cuestión en nuestro campo. El imperativo de sincronización entre diálogos originales y subtítulos, la necesidad de mantener un ritmo de lectura adecuado, la aparente obligatoriedad de usar un máximo de dos líneas en el subtítulo, y las diferentes limitaciones que de ello se derivan, parecían confirmar esa imposibilidad de ofrecer información extra junto con la traducción. Una vez

más, el subtitulado para DVD parece estar abriendo nuevas avenidas y rompiendo con viejos tabúes.

The translator's presence and interference in film texts by use of meta-textual intervention, such as footnotes or glossaries, has always been out of the question in our field. The imperative of maintaining synchronisation between the original dialogue and subtitles, the necessity to maintain an appropriate reading rhythm, the apparent obligation of using a maximum of two lines in the subtitle and the different limitations stemming from this seemed to confirm the impossibility of offering extra information alongside the translation. Once again, DVD subtitling appears to be opening new possibilities and breaking old taboos. [AT]

Díaz Cintas' (ibid) positive viewpoint is an emergent one in the context of the impact of DVD technology on subtitling practices among the more widely shared opinion that the subtitler should aim at being inconspicuous (e.g. Richie 1991, Jaskanen 1999, Castro Roig 2001, Bogucki 2004). Those seeking the translator's invisibility share the opinion that in so far as possible, the viewer should not notice that they are reading subtitles, or have to exert much effort when viewing subtitled AV content. Jaskanen (ibid) uses invisibility as one of the norms by which she determines the quality of subtitles, with more "invisible" subtitles being judged as higher quality, even referring in the context of translation to the invisibility being "violated". Gottlieb (1994 p268) too, highlights the importance of subtitles not being so noticeable so that the "illusion of the translation as the *alter ego* of the original is broken". Issues such as readability come into play when considering the visibility of a translation, as well as fidelity to the source text.

A subtitle full of unfamiliar words, or words the viewer is not used to seeing written, will take more time and effort to read (see Section 6.4.5.1). To replace these with more familiar terms may lower the fidelity of the text, while keeping them may place higher cognitive strain on the viewer. One case where the latter strategy was applied is noted by Nornes (1999 p31), when he mentions Donald Richie, the subtitler of Kurosawa's film 乱 (*ran*, *Ran* 1985). *Ran* (ibid) is a Japanese period drama, and Richie (1991 p16) notes that he thought to introduce some "period colour of [his] own" by using some atypical language in the subtitles, such as the exclusion of occasional prepositions to imitate formal court English, resulting in what Nornes (ibid p31) calls "one of the most abusive translations ever undertaken".

The problem faced by the interlingual subtitler is that his/her work is clearly a translation, as the foreign language soundtrack constantly reminds the viewer. The very nature of the subtitling process, producing a translation where the target audience are “quite overtly not being addressed” (House 1981 p189) as the primary audience, makes it difficult for the subtitler to be successful if we are to believe that “...el logro de un buen traductor de subtítulos es pasar inadvertido...” [...*the sign of a good subtitle translator is to go unnoticed...*] [AT] (Castro Roig 2001 p24).

On the other hand, the abusive subtitler need not worry about this, as he or she is able to take on an interventionist role usually reserved for creator/translators, representing the “ultimate statement against the effacement of the translator prevailing in commercial subtitling” (Pérez González 2006 p271).

2.4 Abusive subtitles research and fansubs

A literature search revealed that the majority of research related to abusive subtitles in anime has looked at fan subs, or fan-produced translations (Nornes 1999, Ferrer Simó 2005, Pérez González 2006, Pérez González 2007, Díaz Cintas and Muñoz Sánchez 2006), with only a small body of research where the issue of abusive subtitling practices in commercial DVD is discussed (Díaz Cintas 2005, Caffrey 2007, 2008b). Fansubs display properties of interlingual and intralingual subtitles, with their proliferation increasing in recent years thanks to the widespread increase of access to the internet and web 2.0’s promotion of user-generated content. Also, the online availability of free software to subtitle digital content and merge the subtitles into one single file with the AV content means that the number of people who have access to the tools necessary for subtitling AV content has increased dramatically.

Published research on fansubs has looked at issues such as the debate behind their history and the legality and ethics of their use (Leonard 2004), the production cycle in their creation (Díaz Cintas and Muñoz Sánchez *ibid*), and the categorisation of the abusive translation procedures found in them, the latter of which is discussed in the following section.

2.4.1 Categorisation of abusive procedures in fansubs

Some of the research into abusive translation procedures found in fansubs in TS does not refer to the procedures themselves using the term “abusive”, exemplified by Díaz Cintas and Muñoz Sánchez (ibid) favouring the term “hybrid subtitles”. They provide a list of the main presentational differences between fansubs and professional subtitling, on the basis of Ferrer Simó (ibid, quoted in Díaz Cintas and Muñoz Sánchez ibid p47):

- Use of different fonts throughout the same programme
- Use of colours to identify different actors
- Use of subtitles of more than two lines (up to four lines)
- Use of notes at the top of the screen
- Use of glosses in the body of the subtitles
- The position of subtitles varies on the screen (scenetiming)
- Karaoke subtitling for opening and ending songs
- Adding of information regarding fansubbers
- Translation of opening and closing credits

Pérez González (2007 p71), reduces this list to four items:

- i. Wide range of fonts and type faces used
- ii. Speaker-colour association
- iii. Use of gloss and explanation as “headnote” on top of screen
- iv. Profusion of titling elements positioned outside the traditional bottom of screen area

Although the majority of studies have focused on fan produced subtitles, the percolation of some of the above-listed procedures into commercial subtitles means that limiting the categorisations to the world of amateur subtitles or fansubs may no longer be justifiable. One result of this relegation to the amateur domain is the disregard of abusive subtitles by some on the grounds that “however interesting and innovative, Nornes’ controversial idea has yet to find its place into [sic] full-scale moviemaking” (Bogucki 2004), resulting in an attitude of resignation where we “have to contend [sic] ourselves with imperfect, ‘corrupt’ subtitling” (Bogucki ibid). As the subtitles used in the present study are found on a commercially released DVD, they are not referred to as fansubs or amateur subtitles.

A distinction is made in this study between *translation procedures* and *translation strategies* on the basis of Howell (2005). Translation procedures

refer to the individual methods used to deal with items in a text, such as pop-up gloss, transliterations or the use of slang. Translation strategies refer to the overall approach taken by the translator to the text, and depending on the strategy used by the translator, different translation procedures will more likely be chosen to deal with culturally marked items.

The distinction between translation procedure and translation strategy facilitates the analysis of data on a micro and macro level. At the micro level, the effects of the procedure used are analysed using individual subtitles, while at the macro level, the effects of the strategy are analysed by comparing excerpts with different subtitle sets. As mentioned in Chapter 1, the abusive procedure being investigated in this study is the pop-up gloss, or as it is termed by the distributors of the source text DVD: the *AD Vid-note*.

2.5 An abusive procedure: The AD Vid-note

The option to view the content of the DVD used in the present study, *ぱにぽにダッシュ* (*paniponi dasshu!* Paniponi Dash! 2007), with pop-up gloss is found in the extras menu of the DVD. The viewer can choose to watch the anime with AD Vid-notes, a term used to describe pop-up gloss which is most likely a form of wordplay on the DVD distributor's name, *ADV films*. The company uses pop-up gloss in several of their DVD titles. In *Paniponi dash!* (ibid) the pop-up gloss is displayed on screen in a white box and is used to explain or comment on culturally marked items appearing in each of the four semiotic channels available: visual verbal, visual nonverbal, audio verbal, audio nonverbal (see Figure 7.5). The culturally marked items include visual references to Japanese popular culture, information about the Japanese voice actors, explanations about Japanese terms in the subtitles, etc. The AD Vid-note is similar to the technique used by the North American TV channel VH1 in a music video programme called *Pop-Up Video*, where general knowledge about people or items in music videos was included in balloons superimposed over the video (Cubbison 2005, see example in Figure 2.1). Some commercial DVDs contain other types of pop-up, such as *The Matrix: Special Edition* (1999), where the viewer can open links to multiple viewing angles of a scene whenever a white rabbit appears onscreen, and *Bridget Jones: The Edge of Reason* (2004),

where the viewer can choose to do an interactive quiz while watching the movie.

The term AD Vid-note is accompanied by the trademark logo (™) on both the DVD cover and extras menu of the titles by ADV films. A search in the online database of the United States Patent and Trademark Office revealed that while an application was indeed filed in July 2003 to trademark Vid-notes by ADV, describing them as “interactive software selectable when viewing a video on DVD containing said interactive software for adding information to the viewed video”, the same application was abandoned in January 2005 (ADV Films 2005), although reasons for the abandonment are not cited.



Figure 2.1. Screen capture from Episode 24 of the VH1 programme *Pop-up Video*, showing pop-up gloss commenting on a music video²

Ortabasi (2006) highlights the similarities the use of gloss shares with the linear subtitling model and also likens them to footnotes found in books. However, she argues that the use of gloss may be better suited to books, which unlike anime have “designated space, separate from the main text for such explicative additions” (Ortabasi *ibid* p287).

The rationale for choosing pop-up gloss as the translation procedure to investigate in this study lies in its popularity among amateur subtitlers (Nornes 1999, Ferrer Simó 2005, Díaz Cintas and Muñoz Sánchez 2006), and its receipt

² Image downloaded from http://en.wikipedia.org/Pop-Up_Video [accessed 20 December 2008].

of recognition in recent research, with Bucaria & Chiaro (2007) noting the possibility of using glosses in instances where the viewers could be expected to lack the necessary cultural knowledge to interpret aspects of the AV content correctly. A previous empirical study carried out using questionnaires also revealed it to be the procedure that just under half of the 58 subjects said should be used to deal with unfamiliar items in the visual nonverbal semiotic channel (Caffrey 2008a). The above reasons, as well as its percolation into commercially released DVDs such as *Paniponi dash!* (2007) contributed to the selection of this procedure for focus of the present study.

The menu screen from which the AD Vid-notes can be selected is shown in Figure 2.2. The distributor's choice of the word "insanity" here shows an awareness that the use of pop-up glosses may result in a somewhat hectic viewing experience for the viewer. However, the negative aspect is neutralised by describing the explanations and pointers as "helpful" to the viewer.



Figure 2.2. Extras menu on *Paniponi dash!* (ibid) where pop-up gloss can be selected

2.6 Translating the visual in AVT

While the pop-up glosses used in the *Paniponi dash!* (2007) DVD provide information about items in each of the semiotic channels, the present study will only be concerned with pop-up glosses that refer to the visual nonverbal semiotic channel. There are several studies that look at the translation of nonverbal items in AVT in both dubbing (e.g. Chaume Varela 1997, Bucaria

2005) and subtitling (Pettit 2004, 2007, Pedersen 2005, 2007, Ortabasi 2006, Valentini 2006, Caffrey 2008a). However, much of this research is limited to the analysis of instances where the visual semiotic channel is referred to verbally in the AV content. Ortabasi (ibid p280) notes that the visual aspect of AV content is often regarded as a secondary source of narrative information by film audiences, and that:

While they constantly deal with the complex relationship between the verbal and visual channels of film, AVT translators have generally shared this prejudice by focussing almost exclusively on translating this “textual information”.

Empirical research carried out by Valentini (2006) seems to support this stance, with the results of a corpus-based study suggesting that visuals only tended to be “translated” when the original dialogue was explicitly linked to them.

Extra leeway is often granted to the dubber in terms of altering the spoken dialogue to fit the visual elements of the scene, such as changing the dialogue to improve the synchrony of the voice dubbing with lip movements of the actors. However, a similar freedom is granted to subtitlers of languages that are not generally spoken by the target culture. Anime produced for a domestic Japanese TV audience can provide difficulties for a Western audience unfamiliar with some of the culturally marked information that appears abundantly in both visual and audio channels of the texts, and the freedom granted the subtitler in some cases allows them to alter the text accordingly.

There is a group of AVT scholars who are looking to the field of multimodal³ analysis in an effort to take the visual aspects of AV content more into account (e.g. Taylor 2003, Pérez González 2007, Szarkowska 2007, Baumgarten 2008) in their analysis of translation strategies used to deal with AV content.

Baumgarten (ibid) argues that AV content has two informational layers, the visual information layer and the verbal information layer. She notes two instances where linguistic reference is made to extralinguistic elements of AV

³ Multimodal texts are described by Kress and Van Leeuwen (1996 p183) as texts in which the meanings “are realised through more than one semiotic code”.

content, i.) to make an object in the extralinguistic context the subject matter of discourse, and ii.) to single out elements for the viewers' attention. By making linguistic reference to an extralinguistic element in the image of AV content, visual-verbal cohesion is accomplished. However, Baumgarten (ibid) notes that references need not be explicit, and that the mere co-presence of visual information and linguistic structures adds to the meaning of both. When pop-up gloss is used, the additional verbal information provided accomplishes one form of cohesion, linking the gloss to the image, and resulting in the provision of extra information about the image.

2.7 Abusive subtitles on commercial TV anime DVD

There are several reasons why abusive subtitling strategies may have become more acceptable and visible in commercially released TV anime DVDs. Three influential factors in deciding the language policies used by distributors are considered in the following: i.) technological factors, ii.) distributor-centred factors and iii.) fan-centred factors. They are discussed in terms of their influence on subtitling practices for TV anime on DVD.

2.7.1 Technological factors

The advent of DVD technology has had a significant impact on the translation of AV content. One DVD can store up to 17 GB, and in theory 32 subtitle tracks and 8 audio tracks can be stored on one single disc (O'Hagan 2007). This increased level of storage means that more language options are available on DVD when compared with video, which O'Hagan (ibid p160) refers to as "personalisation functionality", as "individual viewers can activate certain linguistic support features according to their own language requirements". This includes subtitles for the Deaf and Hard of Hearing, and audio description for the Blind and Visually Impaired. It also allows for the addition of extra features such as interactive minigames, multiple camera angles, directors' commentaries and other trivia. Cubbison (2005 p51) outlines two further changes that DVD technology has brought, especially to the anime market:

The impact of DVD on anime is really based on its ability to do two things: turn subtitles on and off and switch between multiple audio tracks. This capacity of the format rendered the sub versus dub debate moot by providing a product both factions would buy and feel satisfied by the authenticity of their viewing experience. The ability to produce a single DVD

rather than two sets of videotapes cut production and marketing costs, leading North American anime distributors to move quickly to drop their VHS lines.

The DVD format is suited to some abusive subtitle strategies such as pop-up gloss because:

...comparado con el cine o la television, el DVD permite un mayor control por parte del espectador que puede detener la proyección, o retroceder unos segundos, cuando lo crea necesario (Díaz Cintas 2005 p171).

...when compared with cinema or television, DVD gives the viewer more control, allowing them to pause the image, or rewind it for a few seconds when they feel it is necessary. [AT]

This ability to pause the image turned out to be a limiting factor in the experiment design used in the present study, and is discussed further in Section 7.5.3.1.

Another issue relating to technology that influences the acceptability of abusive strategies are the emergent trends in AV literacy brought about by the widespread use of the internet. People are becoming more accustomed to various text, image and sound combinations appearing on screen simultaneously, and may therefore be more receptive to experimental subtitle layouts and abusive procedures in AV content.

2.7.2 Distributor-centred factors

Time is a major factor for distributors of AV content. O'Hagan (2007) points out that the translation of DVDs is showing an inclination towards the *simship* (simultaneous shipment) model of localisation, with global simultaneous launches of localised versions alongside the original. Industrial commentators note that this has led to a situation where "the price wars are fierce, the time-to-market short, the fears of piracy rampant" (Carroll 2004). The time factor may be particularly significant for anime DVD distributors, where the dedicated fanbase means that fansubs of anime are often available for download online long before their official commercial release on DVD. However, the same fanbase may also be a key factor in the motivation to include extras, such as pop-up gloss, providing an incentive (in the form of new content) for the fan to

purchase the DVD, rather than watching a freely available fan subtitled version on the internet.

The anime DVD used in the present study is distributed by the company *ADV films*. *ADV films* began as *A.D. Vision*, the founders of which ran an animation club⁴ in 1992. It is the largest distributor of anime in North America and Europe, and one of the founders has twice been named one of “The 100 Most Powerful People in Genre Entertainment” (Leonard 2004). The fact that *ADV films* is the market leader makes their experimentation with subtitle strategies more significant (taking into consideration the market fears noted above). Two possible reasons for this are: i.) that *ADV films* had its start-up within the fan market and ii.) it was highly dependent on fan purchases for its initial sales, so without the fan network “ADV would have had no market base whatsoever” (Leonard *ibid* p24). This early dependence on and close links to the fan market may be a contributing factor in *ADV film*’s response to meeting the desires of anime fans, which are discussed in the following section.

2.7.3 Fan-centred factors

Technology and distribution factors that may contribute to the inclusion of abusive subtitle strategies in TV anime on DVD have been described above. However, without fan-driven sales, both other factors could be inconsequential. Cubbison (2005 p48) cites an employee of *ADV films* as saying:

many consumers start out as renters and become collectors...We get fan letters from consumers who tell us their personal collections include anywhere from 300 to 500 titles.

This suggests, albeit anecdotally, that a large part of the anime DVD market is based on repeated custom from fans, as well as one-off purchases from multiple customers. Díaz Cintas (2005 p171) suggests that the reason for the acceptability of on-screen gloss and other experimental methods when subtitling less mainstream AV products, such as anime, is that:

⁴Anime clubs began to appear in North America in the 1970s and provided fans of anime with a platform to attend viewings of Japanese animation (Leonard *ibid*).

El perfil del consumidor de estos productos es, por lo tanto, diferente al del espectador medio, y es esta concepción primaria del receptor la que justificaría el uso de nota metatextuales.

The consumer profile for these products is, for the most part, different to that of the average viewer. It is this difference in the target audience that justifies the use of meta-textual notes. [AT]

The profile of these fan/consumers and their attitudes towards text authenticity is considered by Cubbison (ibid), where she argues that while anime fans do regard the anime itself critically as a text, they also scrutinise the work of the producers, distributors and broadcasters of the anime. She sees this as an effort on their part to “safeguard particular anime texts according to varying definitions of authenticity” (Cubbison 2005 p47). Quoting from one anime fan’s entry containing a list of criteria for DVD releases from an internet forum, she lists some of the issues which are important to fans. These include subtitle timing and translation, editing of the video, and strategies for dealing with culturally marked items. In the case of the latter, the fan advises, “do not ‘localise’ the translation by inserting US cultural elements; if we wanted to watch cartoons written for the US market, we wouldn’t be buying anime DVDs” (Musbach 2004, quoted in Cubbison ibid p 47).

This statement represents a call for a foreignising strategy to be used when translating culturally marked items. For many fans of anime, the fact that it comes from a foreign country and that it has this context is a vital part of the viewing experience (Price 2001). The pop-up gloss is one solution for retaining culturally marked items in the translation, allowing for both domesticating and foreignising translation strategies (Venuti 1995) to be used simultaneously. For example, culturally marked items in the dialogue may be transferred in the subtitle, and explained with a pop-up gloss. Also, the use of this procedure in many fansubs would have provided a positive indicator for the distributor that it would be an acceptable, perhaps even desirable, feature for the audience.

That being said, there would be no point in using a translation procedure to which the audience cannot accustom themselves or that they cannot view properly. It can thus be argued the present study needed to be based on empirical evidence to investigate how abusive procedures actually affect the perception of AV content.

2.8 Summary

In this chapter, a definition was provided for abusive subtitles, describing them as subtitles that experiment with and push the traditional constraints present in subtitling. Some of these constraints were discussed, with particular focus on the translator's visibility. A distinction was made between translation procedures and translation strategies, to facilitate later data analysis. The abusive procedure that is the focus of this thesis, the pop-up gloss, was described, and some possible reasons for the inclusion of abusive procedures on commercially released TV anime DVDs were discussed.

The pop-up gloss is an example of a complex abusive translation procedure in subtitling, and its addition to DVD requires significant time, effort and expense on behalf of the distributor. By providing a methodology for studying its effects on perception a foundation will be laid for studies on other subtitling procedures, both ones that are currently being used and those that may be adopted in the future. Furthermore, investigation of the effects that the use of pop-up gloss has on the perception of AV content will contribute to the discourse on the use of pop-up gloss in terms of cognitive benefits and strain for the viewer. The following chapter takes a closer look at the type of AV content that was used in the present study, TV anime on DVD.

CHAPTER 3. TV ANIME ON DVD

3.1 TV anime: A definition

One National Geographic journalist refers to anime as having an “elastic definition” (Lovegren 2003). The reason for this can be seen when the variety of meanings people attribute to the term is considered. The definitions range from the general “Japanese animation” (Cubbison 2005, Howell 2005, O’Hagan 2006a, Pérez González 2007), to more restrictive definitions, such as when it is used to refer to “Japanese animated films” (Lovegren *ibid*, Natsume 2000). There are also content-based definitions of anime, and its print-partner, manga⁵, exemplified by Phillips’ (1996 p193 *emphasis in original*) definition:

Es sind Erzählungen, die – obwohl sie sich über mehrere tausend Seiten erstrecken – keine nennenswerten Inhalte aufweisen, realisiert in einem besonderen Zeichenstil, der sich vor allem durch starke Wechsel zwischen dynamischen, *action*-geladenen und statischen Sequenzen und eine grenzenlos übertriebene Mimik und Gestik der Figuren auszeichnet.

They are stories, that – although they span thousands of pages – display no particular content, realised in a particular drawing style, which can be distinguished primarily by the stark changes between dynamic, action-packed sequences, and static sequences, as well as the boundlessly exaggerated movements and gestures of the characters. [AT]

Napier’s (2001 p10) description of anime also focuses on content, when she describes it as:

A medium in which distinctive visual elements combine with an array of generic, thematic and philosophical structures to produce a unique aesthetic world.

Price (2001 p154) criticises such content-based descriptions, claiming that:

...the most common mistake that people make is in categorising it as a style of animation...Actually, the only thing that classifies *anime* as, well, *anime*, is the fact that it is made in Japan by Japanese artists within a Japanese context.

This definition, based on the production process rather than the content, is adapted for the present study, where the term TV anime will be used to refer to animation produced in Japan primarily for a domestic Japanese TV audience.

⁵ Manga are comics made in Japan, which tend to be printed on low-grade paper and are published primarily in black and white (Lankshear and Knobel 2006).

3.2 Anime industry and production processes

The anime industry and many of the production and distribution processes that drive it would be impossible without the worldwide popularity enjoyed by anime and manga. Japan produces around 60 % of worldwide animation (Shūji 2006) and reports from the Japan External Trade Organisation (JETRO) indicate that the market for content based on anime, which includes videogames and character merchandising, was estimated to be worth around €13 billion in 2004 (JETRO 2005 p2). Despite the impact of translation on the genre's success, there is a paucity of research into audience perception of translated anime.

Covering a large range of genres, the anime market can be divided broadly into three categories: feature-length films, TV shows, and video/DVD releases. A survey carried out by the company Video Research revealed growth in the amount of animation broadcast on Japanese television, with a total of 2,850 TV anime programmes broadcast in 2003, up 102 from the previous year and the second consecutive annual rise (JETRO ibid p5).

A typical example of the production cycle for anime (Figure 3.1) shows the variety of financial backers that are involved in the production of anime. These backers include producers of videogames and other merchandise, as well as the original authors of games or manga destined to be developed into anime.

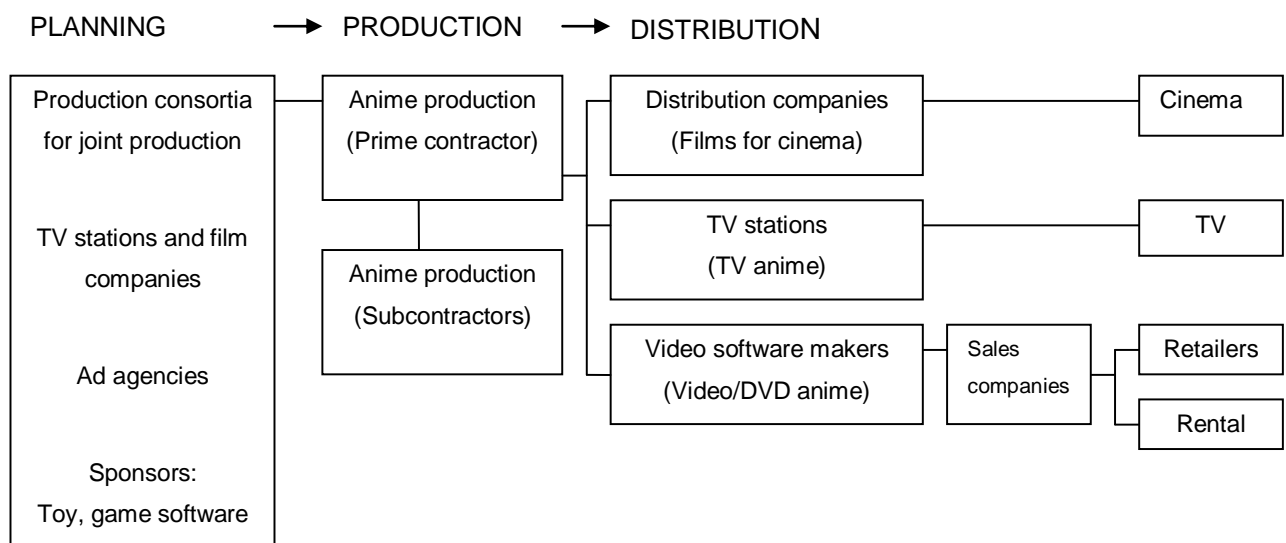


Figure 3.1. Example of the production cycle for anime (adapted from JETRO 2005 p3)

The variety of financial backers is one factor resulting in anime and other media such as videogames and manga sharing themes and pictorial elements (Natsume 2000, Kubo 2004). Multiple financial backers came about as a result of the establishment of production committees in the mid-nineties, allowing funding from a variety of backers to finance one project. In return, different parties gain the rights to a certain aspect of the production, as Zahlten (2007) illustrates:

...so mögen bei einem beliebigen Film die Rechte an der Fernsehausstrahlung bei mehreren Fernsehstationen verbleiben, Tōhō behält den Kinovertrieb, Bandai den DVD Vertrieb, Kadokawa die Verarbeitung als Manga, Geneon den Vertrieb des Albums, Little More den Vertrieb des Buches, eine Talentagentur setzt ihre Schauspieler ein, Avex International erhält die Filmvertriebsrechte für Asien, etc.

...so for any particular film, several TV stations can hold the rights for TV transmission, Tōhō for the cinemas, Bandai the DVDs, Kadokawa the conversion into manga, Geneon the albums, Little More the book, a talent agency uses their actors, Avex International has the film rights for Asia, etc.
[AT]

This was undoubtedly one cause of the recent increase in domestic broadcasting of TV anime within Japan (JETRO 2005). It became cheaper and less risky to invest in the production of anime, a fact that has led to 417 anime films being made available for rental in Japan in 2007, the highest figure since 1971 (Zahlten *ibid*). The change in production processes has also had a knock-on effect on the subject matter of anime, with themes chosen which can be marketed to as many groups as possible in order to increase the likelihood of maximum profitability for the varied backers, although it should be noted that the groups targeted are primarily, if not exclusively, Japanese consumers (Levi 2006).

The effectiveness of the synergy between production, marketing and the variety of distribution channels is made clear by the example provided by Zahlten (*ibid*), in a list of product offshoots of the production committee backed manga, 新世紀エヴァンゲリオン (*shinseiki evangerion*, Neon Genesis Evangelion):

Manga	(February 1995 -)
TV series	(October 1995)
20 videogames	(1996-)
Three films	(1997-1998)
Two further manga series, based partially on videogames	(2003-)
Four pachinko machines	(2004-)
16 music albums	(2005-)
Four new films	(September 2007 -)

The complex and reciprocal relationship between anime and manga is reflected in their shared history. The first major TV anime hit in Japan was 鉄腕アトム (*tetsuwan atomu*, Astro Boy) in 1963, based on the manga of the same title. This sparked a trend of creating TV anime based on manga, as well as the simultaneous release of both manga and anime (Natsume 2000). Kubo (2004) discusses the results of a survey including 87 TV anime, noting that:

視聴率が最も高い15番組のうち14番組はコミックを原作にしたものです。このように、日本のテレビアニメは質・量ともにコミックにより支えられています。

Of the 15 programmes with the highest viewer ratings, 14 were based on manga. In this way, manga bolsters both the quality and quantity of Japanese TV anime. [AT]

The relationship between anime and manga has resulted in the marked influence on the visuals used in both forms, with visual symbols and icons being shared by both. Price (2001 p154) argues that anime “owes a lot of its characteristics to manga culture and aesthetics”. Two aspects of anime whose origins she links back to manga are the use of enlarged eyes, pioneered by Osamu Tezuka, which she claims grew from the need to show emotion on the two dimensional black and white pages of manga comics, and the use of vibrant hair colours in anime, which she sees as:

...an evolution of character depiction in anime. To clearly differentiate the characters from one another, *manga* artists would have to shade their hair in varying degrees of white, black and grey (Price *ibid* p155).

When anime appeared, these different shadings were changed into different colours. While some visual aspects present in manga and anime such as eye

size and hair colour may have little influence on viewer comprehension, others may present difficulties for the viewer unfamiliar with anime or manga tropes.

3.3 Anime genres

Anime covers a multitude of production categories and genres, ranging from the critically acclaimed feature-length and made-for-TV works of directors such as Kon Satoshi, to TV anime and anime made specifically for DVD release. These genres were defined in order to distinguish anime in terms of their length, content, structure and target audience (Phillips 1996). A limited sample of the genres include, 歴史 (historical), 青年 (male youth), 少年 (boys), 少女 (young girls), 子供 (children), 幼年 (infants), SF (science fiction), and スポーツ (sport). As well as the genres listed here, other genres include sexually explicit themes, while others are violent. However, this is just a narrow section of the broad range of anime, which as they are often based on the content of manga, include the same variety of subject matter. However, the more adult-oriented genres traditionally received more attention in the past leading to the misconceptions that anime mainly features “big-boobed femme fatales that gruesomely destroy and decapitate anything in sight” (Price 2001 p159).

The genre used in the present study is a high school comedy anime. Anime with a high school theme is regarded by some as being easier to translate than other genres such as SF or historical (Henry 2006). Admittedly, the dialogue in high school anime tends to be modern, colloquial Japanese, which provides relatively few problems for dialogue translation in terms of register or archaic language, compared, for example, with historical anime. Nevertheless, the genre is replete with culturally marked references, appearing in each of the semiotic channels. Also, the assumption that popular culture texts are easier to translate than more “highbrow” texts is not always accurate, as professional translators are expected to uphold standards independent of the source text type and a popular culture text, including high school comedy anime, “can be just as challenging to its would-be translators as any other genre” (Rovira Esteva 2007 p28).

3.4 Anime and cultural knowledge

The home video industry, from small anime companies to major distributors like Buena Vista Home Entertainment, have been forced to acknowledge anime fans as consumers willing to invest a substantial amount of money in products that meet their demands and scornful of products that don't (Cubbison 2005 p55).

Research into the issues surrounding anime fans and their influence on distributors in North America (Cubbison *ibid*) reveals that the arrival of DVD technology has had a significant effect on distributors of anime, mainly in terms of the necessity to provide both subtitled and dubbed versions on one disc, and also in terms of the quality of the subtitles provided. On the other hand, North American fans have little influence on Japanese distributors, as Levi (2006 p44) claims that from the perspective of distributors in Japan, "anime and manga are still largely targeted exclusively at a Japanese market without much consideration for the international market it actually enjoys". Price (2001 p156) concurs, claiming that "fans around the world might be surprised to know that *anime* is created with only the Japanese audience in mind". The same opinion is also shared by some creators of anime, with Ledoux (1997 pp31-32, cited in Price *ibid* p156), quoting the Oscar-winning anime director Hayao Miyazaki as having said:

I discovered that my work was a product of Japanese historical perspective and sense of nature... So I have no plans to start making films with a global market in mind... Japan will always remain very much the foundation of my work.

Some see the "inherent Japaneseness" of manga and anime as testimony to their virtues as tools for intercultural education. Kusaka (Kusaka and Koo 2006 p35) likens the use of anime as a means of cultural exchange to Napoleon's use of priests to teach the Corsicans about the French ideals of "liberty, equality and fraternity" by having them imbibe the ideals on the "spoon" of religion, declaring:

And now Japan has a pair of spoons of its own, manga and anime, from which people around the world are imbibing Japanese ideas...I believe that manga and anime can serve as the spoons for transmitting Japan's culture and aesthetic sensibilities to the rest of the world...

Others echo these thoughts, foregrounding the cultural aspects of anime, and claiming that certain series⁶ can even be regarded as “a manual for proper Japanese social behaviour” (Schilling 1997 p224). The notion of anime as a source of education about Japanese culture recurs in Price’s (2001 p169) defence of hardcore anime fans as “diligent intercultural diplomats who are bridging the gaps between East and West, all for the sake of entertainment”. However, in order to verify anime’s applicability as a didactic tool, it would be advantageous to determine the level of cultural awareness being reached by viewers of anime. TV anime are largely created with the domestic Japanese TV audience in mind, and this entails a large amount of culturally marked content.

The anime industry in North America is primarily an import-based industry, rather than a Japanese export industry, a fact that is highlighted by the unusual way in which anime’s popularity grew in North America. Leonard (2004 p4) describes how fans began by importing anime through free distribution fan networks, even years before the arrival of the internet, and how the dissemination process “flouts theories of globalization as directed by American cultural imperialism, for Americans ‘pulled’ Japanese cultural products to America *en masse* without coercion by Japanese industry”. As such, anime can be seen as a market created by fans, and one that is “still very often catered to by companies started by fans” (Levi 2006 p46).

One possible drawback of content so rich in culturally marked items is that it can be difficult for subtitlers to translate the vast quantity of culturally marked material they encounter within the spatio-temporal confines imposed by subtitling. In a study where she attempts to determine the level of understanding of the culturally marked aspects in anime, Levi (*ibid*) looks to fanfics as a means of determining intercultural abilities of the fans that create them. Fanfic, or fan fiction, is the name given to stories created and disseminated by fans of fictional content using characters from the content. Schultz (2008) suggests that the popularity of the internet in the 1990s brought

⁶ Schilling (*ibid*) uses the TV series さざえさん (*sazaesan*, Sazae-san) by Machiko Hasegawa as his example. It has been broadcast on the Japanese TV channel Fuji Television since 1969.

fanfics online, replacing fanzines or fan magazines. Levi (2006) raises concerns that most of the fanfics written by North Americans reveal quite a low level of intercultural competence. Another motivation for an empirical study into audience perception of translated anime becomes apparent in this light, as it will add to the discourse on anime as a cultural product and determine to what degree, if at all, the culturally marked items are being translated and understood.

3.5 The translation of TV anime

While much research into anime has been carried out in film studies (Mc Carthy 1999, Nornes 1999, Brown 2006) and from the perspectives of intercultural studies (Napier 2001, 2005, Cubbison 2005, Manzenreiter 2007, Schodt 2007); there is a paucity of research into anime in the field of TS (O'Hagan 2006a). This includes the notable lack of empirical studies into the perception of both commercially and fan subtitled anime produced using abusive subtitling strategies. The importance of translation for the currently popular image of Japan seen through manga and anime is not lost on researchers in the field of intercultural studies (Yoshioka 2007):

つまり日本文化の“coolness”とは、日本文化の内的な本質によるものではなくて、それが異文化の中に移植されること、「翻訳」というプロセスのなかから発生しているように思えます。

The “coolness” of Japanese culture is not an intrinsic aspect of Japanese culture, it is something that is born from its transfer into another culture, from the translation process. [AT]

Considering the crucial role played by translation in its worldwide popularity, the paucity of study in this area is surprising, but can be accounted for in some way by the lack of linguistic competence on the part of TS scholars (O'Hagan *ibid*).

In a special issue of the journal *Perspectives: Studies in Translatology* dedicated to anime, manga and video game translation, O'Hagan (*ibid*) gives an overview of the current state of research on anime in TS. In the following a review of some of the relevant research that has been carried out to date in the translation of anime and manga is presented under the following headings:

- Transfer of character voice in dubbed and subtitled anime
- Transfer of humour in dubbed anime
- Dealing with pictorial elements

3.5.1 Transfer of character voice in dubbed and subtitled anime

In his doctoral thesis, Howell (2005) takes a descriptive-stylistic approach to analysing translation strategies and procedures used in subtitling and dubbing anime dialogue. The source texts chosen for his study differ from the translations in the subtitle set of the present study because his source texts did not contain any elements he considers to be “abusive” in the style described by Nornes (1999) and did not contain radically foreignising strategies. As mentioned in Chapter 2 (see Section 2.4.1), Howell distinguishes between textual procedures and intercultural strategies adopted by the translator. Studying the strategy used in several texts, he goes on to draw some conclusions on conventions used in the translation of anime dialogue.

In a later paper based on the same thesis, Howell (2006) concludes that procedures such as dialect substitution, autonomous insertion of swearing, humour or foreign language cannot be found in anime subtitling and that subtitling can therefore be considered as a foreignising translation strategy. Howell’s research provides an insight into the translation process involved when translating dialogue. While the topic of his research is translation strategies, it is concerned primarily with dialogue translation and as such is not directly relevant to the current study of the perception of subtitled TV anime. However, his distinction between translation strategies and procedures is adopted in the current study (see Section 2.4.1).

3.5.2 Transfer of word play in anime

Mas López’s (2004) article on the dubbing of TV anime into Catalan is also concerned with dialogue translation and it provides some examples of the changes made in the translation process in order to “give the translated version the natural phrasing and appropriateness of the target language”. One example, from an episode of the TV anime *クレヨンしんちゃん* (*kureyon shinchan*, Crayon Shinchan), is based on a pun using two similar sounding words in Japanese. Shinchan, the protagonist, annoys a ceramic artist, Mr.

Ōkama, by calling him a transvestite (お釜 *okama*) all the time, due to the slight difference in the pronunciation of the first vowel. In the Catalan version the pun was altered to the confusion between the words “workshop” (taller) and “straw loft” (paller), a similar play on words with one letter difference between the words, resulting in the artist being annoyed that Shinchon is looking for cows in his workshop.

Price (2001 p164) provides another example where complete substitution is used, this time in the dubbed version of the series ふしぎ遊戯 (*fushigi yūgi*, *Fushigi Yugi*), in a scene where a character says they found a great stick earlier in the day. There is a play on the word すてき (*suteki*), which can be used to mean both “great” and “stick”. This is translated as “I found a great stick today, CANE you believe it?” in the fansubbed version, while the official dubbed version of the same sentence is “What do you call a sumo wrestler coming out of the closet? A power outage.” These two versions take a very different approach to the translation problem presented; with the first creating a new joke to replace the original in the source text, while the second completely substitutes the original dialogue.

The discrepancies between the fansubbed and dubbed versions also highlights the different priorities present in each translation type, with the fansubbed version attempting to retain the original meaning, and the dubbed version arguably giving more priority to retaining the original effect. It also serves as an example of the different constraints involved in dubbing and subtitling, with the creator of the dubbed version generally having more freedom to alter the dialogue as long as lip-synch is maintained. Furthermore, some anime have relatively simple animation for mouth movements, which makes lip synchrony easier to achieve. However, in the case of less widely spoken languages like Japanese, the difference in freedom granted to subtitlers and dubbers is not so large, as the subtitler may also be free to make more changes to the dialogue.

3.5.3 Dealing with pictorial elements

Issues concerning the translation of pictorial elements tend to have been researched more from the perspective of manga (Phillips 1996, Jüngst 2004, O’Hagan 2008a). However, many issues involving pictorial elements are

shared by both media. While the issue of flopping⁷ pages only concerns manga, it provides us with an example of the details in the image that may go unnoticed by those unfamiliar with Japanese culture (Price 2001 p165):

Tradition dictates that Japanese should always wear the left side of the kimono over the right side. The only time people ever wear the right over the left is when they are dead and ready to be buried. So in effect, flopped *manga* constantly features images of corpses walking around its pages.

Flopping has become less of an issue in manga, as the trend is now to retain the Japanese reading direction in commercially published manga translations (Jüngst 2004), but there are other aspects of the image that the viewer may be unfamiliar with, and can in turn present difficulties for the translator. Phillips (1996 p196) lists some of the items that appear in manga, and anime, that are not used in Western comics. These include the use of a bubble on the nose of a character to indicate that they are sleeping, and a nosebleed to indicate sexual excitement. Jüngst (ibid p97) notes that these elements tend to remain unchanged in anime, no matter how foreign they may look, because “as in Japonism, the different visual character is part of the attraction”.

While the image often remains unchanged, the same cannot be said for the dialogue that refers to the image. Dialogue may be changed for censorship reasons, such as the example cited by Natsume (2000 p3):

When *Maison Ikkoku* was aired in France, scenes of adults getting drunk were changed to show the adults drinking lemonade instead of alcohol, resulting in the sight of peculiar Japanese getting drunk on lemonade!

Other reasons for changing the dialogue can be that the distributor wants to cover up the Japanese origins of the text, or to avoid using words that are unfamiliar to the viewer. Price (ibid pp163-164) highlights an example of this when she remarks, “in dubbed *anime*, whenever a character is eating *omochi*, *onigiri*, *dango*⁸, or anything remotely round in shape, it is usually translated as a doughnut”.

⁷ Flopping refers to the procedure of reversing the pages of manga comics to suit the western reading direction, resulting in several changes to the original image, including many characters becoming left-handed.

⁸ *Omochi*, *onigiri* and *dango* are all rice-based Japanese food items.

The dialogue referring to the image may also be changed in order to keep the humour of the content intact. The Catalan translation of the anime *Crayon Shinchan* illustrates this in an example where Shinchan's friend has a head the shape of an *onigiri*, which in the Catalan version is said to resemble a billiard ball (Mas López 2004).

Another aspect of the image quite particular to the manga and anime genre is the use of onomatopoeia and mimetic expressions, which O'Hagan (2006b) describes as "a distinctive linguistic device in the Japanese language to create an impact similar to sound effect in cinema". They are incorporated in the image and often drawn in such a way that they reflect their meaning. The example in Figure 3.2 comes from the anime *Doki Doki School Hours* (2005) and shows the mimetic expression *ぴよんこぴよんこ* (*pyonko pyonko*), which refers to a child performing a hopping action. In the scene depicted in Figure 3.2, the Japanese characters in the expression jump up and down on the screen, further integrating the meaning of text and image. These expressions present particular difficulties for the translator:

Gerade bei den Onomatopoeika, den Gestaltungselementen, die nicht eindeutig als verbale oder als piktorale Teile zu klassifizieren sind – den "Scharnieren" zwischen Text und Bild – werden ganz unterschiedliche Realisierungsformen in den ZS-Manga gewählt (Phillips 1996 p196).

Particularly onomatopoeia, the compositional feature that cannot be clearly categorised as verbal or pictorial – the "hinges" between text and image – are dealt with in many different ways in the target text manga. [AT]

O'Hagan (ibid) discusses the decisions made by manga publishers relating to these expressions, noting that some "translate all onomatopoeic expressions...whereas some other publishers use the strategy of 'transfer' of such expressions whereby leaving them untranslated". The translator's note to the manga *Japan Inc. An introduction to Japanese Economics* (Ishinomori 1986/1988) claims the reason for leaving these elements untranslated is that "since their function is partially decorative, they can be considered part of the artwork, or enterprising readers can devise their own translations".



Figure 3.2. Example of mimetic expression in TV anime, original expression behind subtitle surrounded by yellow box, translation surrounded by red box (*Doki Doki School Hours* 2005)

Dealing with these expressions is also an issue when subtitling TV anime, as a variety of translation procedures can be used to deal with onomatopoeia and mimetic expressions, including the use of captioning on the top of the screen, or the translated text being placed directly over the expression (see Figure 3.3). Both procedures have advantages and disadvantages, with the latter integrating the translation more into the AV content, while the former allows for a more unobstructed view of the original expression.



Figure 3.3. Two different procedures for translating mimetic expressions (top *Doki Doki School Hours* 2005, bottom *Paniponi Dash!* 2007)

Ortabasi (2006) looks at AVT practices from a film studies perspective and cites Kon Satoshi's anime film *千年女優* (*sennen joyū*, Millennium Actress 2001) as an example of a film with a highly visual narrative with many culturally marked references in the visual nonverbal channel. She notes that the prominence of cultural information in the visual nonverbal channel may be a factor in such a film not being a commercially successful export, because it "privileges image over narrative, remains mostly 'untranslated' and therefore misunderstood" (Ortabasi *ibid* p280).

3.6 TV anime as a source text

Much of the English-language research into anime in AVT follows the general trend in AVT to focus on feature-length or critically acclaimed texts, ignoring programmes that “enjoy lesser cultural prestige” (Díaz Cintas 2004b). All of the articles in the volume edited by O’Hagan (2006a), and most of the research into anime in general, focus on anime or anime directors that are considered canonical in Europe and North America. This leaves a significant gap in the research into the translation of TV anime, which, as outlined above, makes up a large proportion of the anime produced yearly. The focus of the present study on abusive procedures makes it more suitable to use a non-canonical text that allows distributors to be more experimental with their subtitling, something that may be unacceptable in more critically-acclaimed feature length anime destined for a broader, non-specialist target audience.

TV anime constitutes a source text that is aimed almost exclusively at a Japanese domestic audience. While the claims of Miyazaki above suggest that he does not plan to make movies with a foreign audience in mind, it should be highlighted that Miyazaki is in a position to do so, as he is already established on an international stage. It is questionable to what degree other moviemakers share his outlook. In the case of TV anime, however, the market is more clearly defined (i.e. made for domestic Japanese consumption), and for this reason it is expected to contain more culturally marked content than its cinema-bound counterpart.

In addition, TV anime seems to be a favourite source text for fan translations, possibly because of the lower availability of commercial translations for TV anime. TV anime could represent a more authentic form of Japanese cultural products for fans when compared with feature-length anime, as the former are generally produced solely with a domestic Japanese audience in mind.

3.7 Perception of TV anime on DVD

When deciding which aspects of TV anime to investigate, research from the field of intercultural studies provided a useful point of departure for the present study. Jungheim (2004, 2006) has carried out several studies investigating viewer perception of Japanese refusal gestures. In one of these studies (Jungheim 2004), he compares the understanding of refusal gestures in video excerpts of real people in role-play situations and animated excerpts from the series *Sazae-san*. His reason for comparing real-life and animated excerpts was that previous research (Jungheim 2000) revealed that refusal gestures in the anime *Sazae-san* were performed for a longer length of time and in a more exaggerated manner than those performed by native speakers in role-play situations. Despite the longer and more pronounced performance of the gestures in the animated excerpts, the results indicated that both Japanese native speaker and Japanese learners found it more difficult to interpret gestures in animated excerpts. Jungheim (2004) attributes this to the increased complexity of the animated excerpts, when compared to the real-life excerpts⁹. The results of the study provided the initial motivation to investigate viewer perception of nonverbal items in subtitled anime, and the effects that translation can have on it.

In carrying out a study into the perception of a text type like TV anime, which often has semantic items appearing in multiple semiotic channels simultaneously, it would be unfeasible to focus on all aspects of the AV content. Figure 3.4 illustrates the semantic complexity often present in anime in an example from *Paniponi dash!* (2007) in which the dialogue is translated by a subtitle at the bottom of the screen, and graphical text in the top-right corner is translated by text placed near to the original Japanese text. The sheer volume of information in the semiotic mix made an in-depth analysis of a limited range of content from the anime text preferable, and as such two aspects were selected: the visual (nonverbal) content and the subtitles.

⁹ All real-life excerpts were recorded in a classroom with a desk between the interlocutors, while animated excerpts were taken from a variety of contexts, such as at home, in public baths, etc.



Figure 3.4. Screen capture illustrating part of the semiotic mix present in TV anime - note the Japanese text in the original and its English translation in the upper right corner of the image (*Paniponi Dash!* 2007)

Visual nonverbal content and subtitles were chosen as they belong to the top two semiotic channels in subtitled media ranked by Gottlieb (2005a) as having combined approximately 72% of the total informative impact. This percentage may even be revised upwards when AV content in a lesser-known language such as Japanese is investigated, because viewers will be more dependent on the subtitles (and less on the audio verbal channel) to understand the dialogue. Subtitles are generally the translator's only means of mediating the information that is available in all channels of the AV content and are therefore a valuable aid in any study of the perception of the meaning of subtitled AV content. As anime per se is considered a form of visual media, the focus was narrowed down to one aspect of the visual nonverbal content, culturally marked visual nonverbal cues.

3.7.1 Culturally marked visual nonverbal cues: Introduction

The visual content of TV anime is one of the features that distinguish it from animation made in less "exotic" locations, and the genre contains an abundance of culturally marked items in the visual nonverbal semiotic channel. Chaume Varela (1997 p324) notes that while culturally marked items often appear in AV content, with Western cultures sharing many of them, cultures that are more distant may have difficulties understanding "certain signs, scripts or schemata". Viewers unfamiliar with Japanese culture may have difficulties understanding

visual elements of TV anime, particularly those items which are culturally marked. For the present study a definition for the term culturally marked visual nonverbal cue (CVNC) was necessary. Before providing a definition for CVNC, the term will be divided into its two constituent elements: *culturally marked* and *visual nonverbal cue*, which are dealt with separately below.

3.7.1.1 Culturally marked items: definition

After noting the tendency for authors to avoid concrete definitions when discussing terms such as “cultural references”, Aixelá (1996 p58) defines “culturally specific items” as:

Those textually actualised items whose function and connotations in a source text involve a translation problem in their transference to a target text, whenever this problem is a product of the non-existence of the referred item or of its different intertextual status in the cultural system of the readers of the target text.

While this definition is adequate, the use of the term culture specific item will be avoided in this study because while the intended meaning of an item may be specific to the source culture, or those familiar with it, the item itself is not necessarily culturally specific. Take, for example, the narrative technique in anime of a cut to a character sneezing when they are being talked about and they are not aware of it. Although the sight of a character sneezing at such a moment may appear unusual to a viewer unfamiliar with the technique, it is certainly not culturally specific. The connotative meaning, similar to the notion of “my ears are burning”, however, is specific to certain cultures. For this reason, perhaps the term “culturally marked”, as used by Howell (2005) is more suitable because although the items do have a particular imprint of the source culture (in their connotative meaning), they may not necessarily be specific to the culture. For that reason I will adapt Aixelá’s definition for use in this study under the modified label of culturally *marked* item.

3.7.1.2 Visual nonverbal cues: definition

The broad range of meaning involved in the concept of nonverbal communication (c.f. Wierzbicka 1995) means that a more restrictive term is required in the framework of the present study. There are several aspects of communication that cross language and culture, such as gesture and idioms,

and these have come to be defined by terms such as “borderline features” (Bucaria and Chiaro 2007) and “culture-lingual short circuits” (Antonini 2005 p214) by researchers writing on the perception of subtitling and dubbing. Here Caffrey’s (2008a p165) definition of the term visual nonverbal cue (VNC) will be adapted, that is, a “nonverbal item appearing in the image of an audiovisual text which has an intended secondary, connotative meaning”.

3.7.1.3 Culturally marked visual nonverbal cue: definition

Aixela’s (1996) definition of culturally specific items and Caffrey’s (ibid) definition of visual nonverbal cues are combined to define the term culturally marked visual nonverbal cue for the purpose of the present study.

A culturally marked visual nonverbal cue (CVNC) is a nonverbal item appearing in the image of an audiovisual text whose function and connotation in the source text involve a translation problem in its transfer to the target text whenever this problem is a product of the non-existence of the item or of its different intertextual status in the cultural system of the target viewers.

This definition is admittedly somewhat similar to Pedersen’s (2005 p2) definition of an extralinguistic culture-bound reference (ECR), which he defines as a:

reference that is attempted by means of any culture-bound linguistic expression, which refers to an extralinguistic entity or process, and which is assumed to have a discourse referent that is identifiable to a relevant audience as this referent is within the encyclopedic knowledge of this audience.

However, the important distinction between CVNCs and ECRs is that Pedersen’s (ibid) ECRs require the extra-linguistic (or nonverbal) element to be referred to by a linguistic expression, which is not the case in the definition of CVNCs. As such, CVNCs include intersemiotic translation, or translation between semiotic channels (Gottlieb 2005a). While traditional subtitles do not usually provide intersemiotic translation, pop-up glosses often do.

3.8 Summary

This chapter focused on the source text genre that was used in the experiment in the present study, TV anime. A description and definition of TV anime have been provided, as well as an outline of relevant research into the translation issues involved in both anime and manga. The paucity of empirical research into the perception of subtitled anime has been highlighted, and the contribution that empirical research on the perception of subtitled anime could make to the discourse on translated TV anime as a didactic tool concerning Japanese culture mentioned. Finally, the rationale was provided for selecting both aspects of the AV content that are focused on in the study, subtitles and CVNCs. In the following chapter the theoretical framework based on Sperber and Wilson's (1986) relevance theory is explained, and its application in the present study is discussed in detail.

CHAPTER 4. **A RELEVANCE THEORETICAL APPROACH TO THE PERCEPTION OF SUBTITLED AV CONTENT**

Chapter 3 explained the motivation behind the selection of subtitles and CVNCs as the aspects of AV content that would be dealt with in this study. This chapter is concerned with the perspective from which perception is considered and the theoretical framework that is applied to its analysis. Based on Sperber and Wilson's (1986) relevance theory, two aspects of perception are considered: *positive cognitive effects* and *processing effort*.

4.1 Relevance theory

Sperber and Wilson's (ibid) relevance theoretical approach to communication suggests that all overt communication relies on a "communicative principle of relevance", which states, "every act of ostensive communication communicates a presumption of its own optimal relevance" (Sperber and Wilson 1995 p260). Once this principle is satisfied, the audience (or the viewer in the case of subtitled TV anime) proceeds to select the most appropriate interpretation from a range of interpretations. A stimulus is considered optimally relevant to an audience if:

- a.) it is relevant enough to be worth the audience's processing effort
- b.) it is the most relevant one compatible with the communicator's abilities and preferences (Wilson and Sperber 2002 p256).

When interpreting a stimulus, the viewer proceeds by following the path of least effort to construct an interpretation of the stimulus, and stopping when the expectations of relevance are satisfied. Yus (2008) highlights two conditions that must be satisfied for the expectations of relevance to be satisfied: a.) the cognitive effects achieved when an interpretation is optimally processed are large, and b.) the effort required to process it optimally is small. This results in a trade-off between the cognitive effects and processing effort where the addressee (or the viewer) will attempt to maximise cognitive effects while minimising processing effort (Gibbs and Tendahl 2006). In turn, subtitling procedures can be investigated in terms of their relevance to the viewer using the following criteria:

Relevance of an input to an individual

- a.) Other things being equal, the greater the positive cognitive effects achieved by processing an input, the greater the relevance of the input to the individual at that time.
- b.) Other things being equal, the greater the processing effort expended, the lower the relevance of the input to the individual at that time (Wilson and Sperber 2002 p252).

As explained above, relevance theory was initially developed to describe communication processes. Relevance theory has been applied in AVT to describe reduction strategies used by subtitlers (Kovačič 1994), the role of the viewer's cognitive environment on their decoding and understanding of subtitles (Kovačič 1995), and subtitling constraints (Bogucki 2004). However, the application of relevance theory in the experiment outlined in this thesis is more similar to Gambier's (2003 p185) suggestion, where he proposes that the effectiveness of a translation could be determined by following the guideline that "the greater the viewers' processing effort, the lower the relevance of a translation".

In the present study, the effects of pop-up gloss on the perception of CVNCs and subtitles in TV anime are determined using both criteria for measuring the relevance of an input to an individual as suggested by Wilson and Sperber (*ibid*): positive cognitive effects and processing effort. The effective use of pop-up gloss should increase the relevance of the excerpts used in the study, with subjects experiencing more positive cognitive effects at a similar or lower level of processing effort. In the following, the terms positive cognitive effects and processing effort are given operational definitions for the present study, and the methods used to measure them are introduced.

4.1.1 Positive cognitive effects

Wilson and Sperber (*ibid* p251) define positive cognitive effects as "a worthwhile difference to the individual's representation of the world – a true conclusion, for example". This distinction allows for the differentiation between cognitive effects in general and positive cognitive effects. While a false conclusion can be considered a cognitive effect, it cannot be considered a positive cognitive effect. Wilson and Sperber (*ibid* p251 *emphasis in original*) argue that the notion of a positive cognitive effect is

necessary to distinguish “between information that merely SEEMS to the individual to be relevant and information that actually IS relevant”.



Figure 4.1. Nosebleed indicating arousal (*Doki Doki School Hours* 2006)

The following example can be used to explain the difference between a positive cognitive effect and general cognitive effects. The scene from the anime *Doki Doki School Hours* in Figure 4.1 depicts a student (Kudō) having a nosebleed after a conversation with another student (Suetake) with whom he is enamoured. The nosebleed occurs after Suetake makes a play on the phrase 墓穴を掘る (*boketsu wo horu*), the equivalent of the saying “to dig your own grave”, by replacing the first character, resulting in the new phrase 〇穴を掘る (*oketsu wo horu*). The new phrase is translated as “to dig a bum”, and is a colloquialism alluding to intercourse. Upon hearing this, Kudō has a nosebleed, which is depicted in the exaggerated manner illustrated in Figure 4.1, showing a flood of blood bursting from his nose. In such a situation the following thoughts could be entertained by the viewer:

- a) Kudō is ill.
- b) Kudō is afraid.
- c) Kudō is shocked.
- d) Kudō is aroused.*

In real-life situations when someone has a nosebleed, a) is normally more likely to be entertained. However, those who are familiar with the use of the

nosebleed to represent arousal in anime will make the assumption in d), achieving positive cognitive effects. While all of the assumptions in a) – d) represent cognitive effects, only d) represents a *positive* cognitive effect, where the belief in the assumption will make a worthwhile addition to the individual's understanding of the situation.

In the present study it is assumed that in order for a subject to correctly answer a question about a subtitle or CVNC, they must experience positive cognitive effects. This is based on the premise that correctly answering a question provides evidence that a worthwhile difference to the individual's representation of the world has come about by their exposure to the AV material. A wrong answer, on the other hand, may represent a difference to the subjects' representation of the world, but not one that is worthwhile.

As such, in the context of the present study a positive cognitive effect is defined as a worthwhile difference to the individual's representation of the world evidenced by that individual's ability to retain and retrieve new information on the basis of some input. Using this definition of positive cognitive effects, it is possible to determine whether the use of pop-up gloss in an excerpt will result in more subjects experiencing positive cognitive effects than subjects watching the same excerpt with no pop-up glosses. To determine the experience of positive cognitive effects, a questionnaire is used, as it provides a practical means of determining whether information was retrieved and retained from excerpts.

4.1.2 Processing effort

Wilson and Sperber (2002 p252) describe processing effort, or mental effort, as “the effort of perception, memory and inference required” to process an input. Intuitively, the greater the effort to process an input, the less rewarding it is, considering that communication is driven “by the desire for optimisation of resources, and one aspect of optimisation is to keep the effort spent to a minimum” (Gutt 1991 p26).

Processing effort is an attribute of the interaction between a stimulus and a person. Two main factors affect processing effort, the way that the

information is presented, and the accessibility of the context of the information (Wilson and Matsui 1998). Both factors may result in viewers who are not used to seeing AV content with pop-up gloss requiring more processing effort to view an excerpt where it is used. Being an abusive procedure, pop-up gloss is presented in an unusual way in the context of subtitled AV content, and the textual reference to items in nonverbal semiotic channels may initially be a source of confusion for viewers when trying to find the context of the information contained in the pop-up gloss.

As it is a hypothetical construct, processing effort cannot be directly observed, “sondern nur indirekt, über ihre Auswirkungen” [*only indirectly, through its effects*] [AT] (Schultheis 2004 p7). Just and Carpenter (1993) highlight the difficulty of measuring processing effort, noting that as an internal process it does not necessarily leave an impression on any output. There are several types of measure that can be used to measure processing effort. O’Donnell and Eggemeier (1986) propose three categories of measure: subjective, performance and physiological. In a study of cognitive load experienced while reading text on computer, Schultheis and Jameson (2004) include an additional category, analytical measures. Three methods were used to measure processing effort in the present study: subjective measures, performance measures and physiological measures. The use of three distinct methods of measurement employing both consecutive and simultaneous data collection made it possible to approach processing effort from several complementary angles.

Below, all four measures are described, along with their suitability for measuring the processing effort experienced when viewing subtitled AV content.

Analytical measures:

Processing effort can be measured using general information about the subjects and the stimulus being measured. For example, the average subtitle speed of excerpts and the regularity with which subjects view subtitled AV content could be used to form a basis for the prediction of the amount of processing effort required. However, the accuracy with which this

measure may be applied is lower than the other measures outlined below and analytical measures were therefore excluded from the current study.

Subjective measures:

Subjective measures often involve the use of questions that ask subjects to rate the processing effort they required. This may be directly requested, for example, by asking subjects how much effort they required to view an excerpt, or indirectly, by inferring a relation between some other value and processing effort. The latter was the case in the current study, where it was assumed that the more processing effort was required to view an excerpt, the higher the number of subjects experiencing the subtitle speed as fast would be. Paas *et al.* (2003) note that subjective measures are the type of measure most often used in cognitive load studies. As Schultheis and Jameson (2004) highlight, memory and consciousness effects may distort subjective measures of effort. For example, the subject's memory of the excerpt as a whole may be affected by one section that he or she was more conscious of, or merely by the presence of pop-up gloss in an excerpt.

However, the collection of data using subjective measures is relatively uncomplicated and in the context of the present study it provided a practical complementary measure of processing effort for the data gathered with the eye tracker.

Performance measures:

Performance measures are used to ascertain the subject's processing effort from his or her overt behaviour. A secondary task may also be introduced, such as pushing a button when a visual or auditory cue is given (e.g. light turned on in room, beeping noise is emitted). Poor performance in this secondary task is taken as an indicator that the primary task involved a high level of processing effort.

O'Brien (2006) uses a performance measure to determine the cognitive effort required when using the computer-aided translation memory software, Trados, measuring the number of source text words per second processed by subjects. However, Gottlieb's (1995) use of a protest button in a study where subjects

were asked to press a button when they saw a certain type of subtitle is an indication that the use of a secondary task may be poorly suited to measuring processing effort while viewing AV content. Performance measures are only appropriate if the activity yields a sufficiently high rate of accurately observable behaviour (Schultheis and Jameson 2004). Reasons the protest button was not considered a suitable means of measuring processing effort in the current study are explained in Section 6.4.2.

The eye tracker is a useful data collection tool for providing performance measures. It allows for the objective, accurate measurement of the amount of time spent in various areas of the screen, which can later be related to variables used in the experiment. The resulting data can be used to infer the effects of these variables on the processing effort experienced by subjects.

Physiological measures:

The final measure of processing effort discussed here are physiological measures. These measures involve the observation of bodily processes and states, and the use of these observations to infer processing effort. There are a variety of physiological measures that have been used to measure processing effort, including blink rate, heart rate, pupil size, and more recently, brain scanning. Nesbit and Hadwin (2006 p830) note the advantages of physiological measures as they can be used to provide analysis of cognitive load often without the need for overt responses from subjects, claiming that pupillometry has “proven to be one of the most precise physiological measures”. Wilson and Eggemeier (1991) argue that while performance-based measures may be influenced by very high or very low workload, in that the subject may not respond to the stimuli required by a second task, or that these tasks may become monotonous for subjects, physiological measures allow for continuous data recording. Another advantage physiological measures have over subjective measures particularly relevant to our study is that overt responses are not necessary in order for simultaneous data to be collected (Kramer 1991), allowing for the establishment of a more realistic viewing setting.

On the other hand, physiological measures can be affected by factors not necessarily related to cognitive processing. Heart rate may be affected by

physical strain related to environmental factors (Kramer *ibid* p330), and several factors are known to affect pupil size, including the level of lighting, and the intake of some drugs and alcohol. In the present study, the complementary use of subjective and performance measures of processing effort is intended to act as a check for the validity of the physiological measure applied.

In the present study the *Tobii 1750* eye tracker was considered to be a suitable tool for gathering data to be used for the physiological measure of processing effort. The eye tracker provides a means of accurately measuring the size of the pupil, while being less intrusive than head mounted eye trackers, which usually restrict the subject's movement.

4.2 Summary

Chapter 4 explained the relevance theoretical framework that is used to analyse perception in the present study. Using relevance theory as a basis for the present empirical study, two aspects of perception are analysed: positive cognitive effects (PCEs) and processing effort. The data collection tools, questionnaire and eye tracker, which were used to measure both aspects of perception were named and justification for their use given.

The following chapter takes a closer look at eye tracking, how it can be used to measure processing effort, and the eye tracking hardware and software used in the present study.

CHAPTER 5. **MEASURING PROCESSING** **EFFORT USING AN EYE TRACKER**

An eye tracker is a device that “monitors and records the movements of an individual’s eyes” (O’Brien 2006 p185). Eye trackers have a variety of applications, which Duchowski (2003) splits into two main categories: *interactive* and *diagnostic*. In an interactive role, the eye tracker provides an input device for applications ranging from enhancing accessibility (e.g. using gaze as a method of input for quadriplegics [Man and Wong 2007]) to controlling resolution in virtual environments (e.g. in a virtual reality world, the area where the gaze rests may be given a higher pixel count than the surrounding area, leading to more efficient use of the graphic resources of the computer). The role that will be relevant in this thesis is the diagnostic one, distinguished by the eye tracker not affecting either the stimulus being used or the subjects taking part in the experiment, although in the discussion on future research (see Section 9.3) one possible interactive application of the eye tracker with subtitled AV content is mentioned.

Unlike interactive studies, in diagnostic studies the stimulus being shown need not react to the viewer’s gaze. Diagnostic studies are often used in the field of marketing and usability (e.g. study carried out by Google and IBM into how pictures influence online reading [Granka and Rodden 2006, Beymer *et al.* 2007]), psychology (e.g. investigating the cognitive processes involved in reading [Rayner 1998]), as well as process-oriented TS (e.g. attempting to identify differences in cognitive effort related to Translation Memory match types using the translation memory tool *Trados* [O’Brien *ibid*]). A further application of eye tracking in process-oriented TS which combines diagnostic and interactive applications of the eye tracker is the EYE-to-IT project¹⁰, which aimed at developing human-computer monitoring and feedback systems for investigating cognition and the translation process.

5.1 Using eye tracking data to measure processing effort

Eye tracking data can be analysed using a top-down or bottom-up approach (Goldberg *et al.* 2002). The top-down approach involves looking at the data

¹⁰ The Eye-to-IT project was undertaken for the period between January 2006 and December 2008 and was funded by the EU FP6 programme (<http://cogs.nbu.bg/eye-to-it/>).

using a cognitive theory or a hypothesis to be tested by experimentation, while the bottom-up approach is based entirely on the analysis of data without any predefined theory relating to the eye movements. The present study takes a top-down approach based on the premise that the presence of pop-up gloss will affect subjects' cognitive processes resulting in a change in their eye movements. There are several measurements that may be used with the data collected from the eye tracker. These can be categorised roughly into three types of measurement: i.) saccade-based measurements; ii.) fixation-based measurements; and iii.) pupillometric measurements.

Saccades are rapid movements of the eye (up to 500 degrees per second) that relocate the fovea¹¹ of the eye to a part of the visual environment (Duchowski 2003). Saccades can be distinguished from three other types of eye movement:

1. Pursuit eye movements occur when the eye follows a moving target and are markedly slower in velocity than saccades. During pursuit movements, saccades are often made to catch up with targets moving quickly across the visual field (White, 1976).
2. Vergence eye movements occur when we move our eyes inward to fixate on a nearby object (Rayner 1998), such as when we bring a pencil up to our nose.
3. Vestibular eye movements occur when the eyes move to compensate for body movements to maintain direction of vision (Rayner *ibid*).

New visual information is not obtained during a saccade as the eyes move so quickly over the stable visual stimulus that only a blur would be perceived. This process of the cognitive blocking of visual information is referred to as saccadic suppression (Griffin 2001). It is not yet known whether cognitive processing activities are suspended altogether during a saccade (Rayner *ibid*), although there are some indications that they may not be entirely suppressed (Yatabe 2006). The uncertainty regarding the relationship between saccades and

¹¹ The fovea of the eye is the area of the retina where the density of photoreceptive cells is most highly concentrated, allowing for the most accurate vision.

cognitive processing was a factor in the decision not to use saccade-based measurements in the present study. Instead the remaining two measurements, fixation-based and pupillometric measurements, were used.

5.1.1 Fixation-based measurements

Duchowski (2003 p48) defines fixations as “eye movements which stabilise the retina over a stationary object of interest”. The term fixation is used to refer to the eye remaining focused on one area for a set amount of time. Rayner (1998) points out the flaw with the term fixation as the eye never stays completely still or fixates directly on one spot. This is because of tiny movement in the eye caused by nystagmus¹², drift and microsaccades, small tremors of the eye. However, this is catered for in the eye tracking analysis by selecting a pixel radius large enough to take the nystagmus into account, and by the 50 Hz sampling (i.e. 50 recordings per second) provided by the eye tracker employed in the present study (Salojärvi *et al.* 2005).

One of the important basic assumptions behind the application of fixation-based measurements in eye tracking is that visual attention is generally linked to where the eye is directed, although this is not always the case (Posner 1980). While we can dissociate our attention from the point where the eye is fixated, most of the time our attention and our eyes are aimed at the same part of the stimulus field (Rayner 1992). Rayner (1998) lists several studies which show that it is more efficient to move our eyes than move our attention (He and Kowler 1992, Schlingensiepen *et al.* 1986) and asserts that in complex information processing tasks such as reading, and in our case, viewing subtitled TV anime, the link between gaze and attention can be assumed to be quite tight. Therefore we can assume that where the eyes fixate is the part of our visual field on which visual attention is generally focused.

Rayner (*ibid* pp389-390) has highlighted that “considerable data have been collected that demonstrate that eye movements are intimately related to the

¹² Nystagmus eye movements are conjugate eye movements, and can be distinguished by a saw-like pattern of movement (Duchowski 2003 p49).

moment-to-moment cognitive processing activities of readers". Krause (1982) makes three assumptions about fixations and cognitive processes:

1. The eyes fixate on data that is currently being processed.
2. Processing time is directly proportionate to the fixation time of an item. This is also noted by Salojärvi *et al.* (2005), who link fixation time with the complexity of the object under inspection.
3. The sequence of cognitive processing follows the eye fixation sequence.

These correspond with three propositions put forward by Fisher *et al.* (1983), namely that:

1. Fixations accumulate in locations judged to contain high semantic or visual information.
2. Fixations are responsible for perception and are generally considered as a reflection of the individual's cognitive strategy.
3. The fixation sequence allows for the encoding, storing and subsequent reconstruction of the images.

5.1.1.1 Applicability of fixation-based measurements with subtitled AV content

There are several examples of research that have been carried out into the perception of subtitles by psychologists using eye trackers, with much of the research being carried out under Géry d'Ydewalle at the Catholic University Leuven (e.g. d'Ydewalle *et al.* 1985, van Rensbergen *et al.* 1986, Gielen 1988a, Gielen 1988b, d'Ydewalle *et al.* 1991, d'Ydewalle and Gielen 1992, Koolstra *et al.* 1999, d'Ydewalle and de Bruycker 2007). Some relevant research in the field of psychology where fixation-based measures have been used to analyse the perception of subtitles is briefly described in the following.

D'Ydewalle and Gielen (*ibid*) note that subjects experience one-line subtitles as disappearing quicker than two-line subtitles, regardless of presentation time. Two-line subtitles were looked at for a proportionately longer amount of time than one-line subtitles and in some cases the latency time¹³ needed to shift the gaze to two-line subtitles was also longer. Possible reasons for the longer latency time and viewing time are suggested by d'Ydewalle *et al.* (1991). One possible cause of longer latency times with two-line subtitles is that viewers

¹³ The latency time is the time between the onset of the subtitle and the first fixation upon it.

generally prefer to look at the image, and the viewer expects that they will need more time for a two-line subtitle. Referred to as the “length-expectation” hypothesis, this supposes that subjects are able to make an estimate of the subtitle length and shape with peripheral vision, and use this to make an estimate of its display time.

Praet *et al.* (1990) tested whether the proportionately longer viewing time of two-line subtitles may be caused by lateral interference. This lateral interference would be caused by the increased difficulty for the eye to fixate on the first word of the first line of the subtitle, and the sweep necessary to return from the end of the first line to the beginning of the second. However, as the manipulation of the distance between lines in the subtitle revealed no effects, lateral interference was no longer considered a plausible reason.

In another test, Praet *et al.* (*ibid*) compared the viewing of subtitles on screen alone and subtitles on screen with pictures and sound (i.e. a normal subtitling condition). They found that in the normal subtitling condition the percentage time spent in the subtitle area was greater when two-line subtitles were on screen than one-line subtitles, while in the condition where only subtitles were on screen, more time was proportionally spent in the subtitle area for one-line than two-line subtitles. D’Ydewalle and de Bruycker (2007 p202) argue that this is because “the information conveyed by one-line subtitles is far more redundant, given the available pictorial information, than is the case for two-line subtitles”.

5.1.1.2 Units of measurement used with fixation-based data in the present study

On the basis of d’Ydewalle and de Bruycker (*ibid*), the present study used several fixation-based measurements which are all considered as indicators of the cognitive processing that is involved while viewing subtitles: percentage gaze time; average fixation duration; percentage skipped subtitles; and word fixation probability. A description of each of these measures and how they were calculated is given in Section 7.2.4.

One of the weaknesses of using the fixation-based measurements listed above is that they focus primarily on the processing of the subtitles, rather than the overall processing of the AV content, which consists of other items in the visual verbal semiotic channel, as well as the visual nonverbal, audio verbal and audio nonverbal channels. It was hoped that the subjective measure of processing effort in the consecutive questionnaire would provide an account of the effort required by subjects to process all semiotic channels during the viewing of each excerpt. This data was further complemented with data gathered using the physiological measure, pupillometry. Pupillometric measurements, like fixation-based measurements, are collected simultaneously. However, unlike the fixation-based measurements used in the present study (which are based only on fixations in the subtitle area), the pupillometric measurements take all semiotic channels of the text into account.

5.1.2 Pupillometric measurements

The use of pupil size as an indicator of cognitive processes dates back to the 1800s, and since the 1960s, there have been many studies into the behaviour of the pupil in cognitive psychology (Steinhauer 2002). The term “pupillometrics” can be used to describe this field of study (Hess 1972). Many pupillometric studies have suggested that there is a link between pupil size and cognitive load, as: “Je höher die mentale Beanspruchung, desto größer ist die Pupille” [*The higher the cognitive demand, the larger the pupil*] [AT] (Schultheis 2004 p32).

Pupil size offers an advantage over some other measures of processing effort such as blink rate, eye movement and heart rate, in that it offers a practically immediate measure of workload (Iqbal *et al.* 2005). Pupil reaction times are estimated between 100 – 200 ms (Beatty 1982) and 600 ms (Kramer 1991), indicating the speed with which the pupil reacts to stimuli.

Many pupillometric studies have investigated reactions to interactive tasks, such as the memorisation of digit sequences (Kahnemann and Beatty 1966), mental multiplication (Ahern and Beatty 1979), aural vigilance (Klingner *et al.* 2008), as well as interpreting and translation (Hyönä *et al.* 1995, O’Brien 2006, 2008). In all of these studies, pupil size was shown to be a reliable indicator of task

difficulty. They can be distinguished from the present study, as each of the above studies involved an interactive task, while in the research reported here pupil measurements were taken while subjects viewed excerpts in various subtitle conditions, with no specific interactive task involved. However, they provide a relevant foundation on which to develop the analysis of pupillometric data, as well as highlighting important aspects in the design of the experiment and some of the difficulties involved when using pupillometric data.

5.1.2.1 Disadvantages of pupillometry

Despite its many advantages, pupillometry is not the perfect solution for measuring cognitive strain, as many confounding factors can influence the pupil readings (c.f. Krüger [2000], who lists 20 factors that can influence the pupil reading). Some of those factors include gender, anxiety, taste, habituation, schizophrenia, arousal and novelty (Janisse 1977). This makes it difficult to rule out the interaction of several factors when using pupillometric data and has resulted in many experiments that use pupillometry controlling for various factors that may influence pupil dilation.

In a study on the translation of words, Hyönä *et al.* (1995) control for factors including light reflex (by using controlled lighting) and movements of the eye (by having subjects fixate on only one area of the screen and using audio stimuli) and use short tasks at discrete distances from each other to minimize the influence of neighbouring tasks on processing. The use of short tasks at discrete distances was introduced because if tasks occur too closely to each other, the pupil reading for one task may be distorted by the reading from the preceding task.

Other factors that have been controlled are subject-related factors that may affect cognition or pupil dilation, such as head injuries, substance abuse and diabetes (Verney *et al.* 2001). Schultheis and Jameson (2004) claim that the vast majority of research where pupil size is used as a measure of cognitive strain has used three of the following five methods to control influences on data gathered in experiments: i.) constant lighting; ii.) avoidance of eye movements;

iii.) use of acoustic stimuli; iv.) use of many similar short tasks and v.) evaluating only mean values across tasks and subjects.

The difficulty of controlling factors in experiments is one plausible reason for the paucity of published research using pupillometry with audiovisual content. An example of one factor which cannot easily be controlled is the light intensity of stimuli, as can be seen when we look at the physics of light. Not only do different light wavelengths stimulate receptors in the eye differentially, so do different intensities (Steinhauer 2002). This means that even stimuli of matching brightness will have a different effect on the pupil depending on the colour hue. Therefore in the case of AV content it is not reasonable to expect the same light intensity throughout.

Nevertheless, there have been several pupillometric studies where not all of the factors mentioned above were controlled, displaying various degrees of success. While O'Brien (2006) had some success in a pilot study using pupil size as a measure of cognitive effort during the translation process with the translation memory tool *Trados*, Schultheis and Jameson (2004) did not find it a reliable measure of cognitive effort when comparing texts of varying levels of difficulty. It is clear that further research is required to establish the usefulness or otherwise of pupil size in measuring cognitive effort in various contexts, a fact that provides further motivation for the use of this measure in the current study.

5.1.2.2 Application of pupillometric measurements in the present study

As pupil responses do not occur reliably for any one particular stimulus for the reasons outlined in the previous section, observations of several pupillary responses must be combined (Beatty and Lucero-Wagner 2000). Pupillometric measurements are generally taken using one of two methods, the trial-aggregated fine-grained method and the time-aggregated coarse method (Klingner *et al.* 2008). The present study is the first time pupillometry is being used with subtitled AV content. As such, both trial-aggregated fine-grained and time-aggregated coarse methods are described in the following and their applicability to the present study using subtitled AV content is discussed.

Trial-aggregated fine-grained method:

This form of measurement tends to be used in research carried out in the field of cognition (Klinger *et al.* 2008) and deals with the unreliability of single pupil responses by aligning and averaging data over several short trials (Beatty and Lucero-Wagner 2000). An example of this form of measurement can be seen in Just and Carpenter (1993 p320) where they compare the maximum pupil diameters of subjects when reading two sentence types, labelled object relative and subject relative. Object relative sentences are sentences that have a centre-embedded relative clause, such as “The man that the child bit admitted the mistake”. These sentences are considered to be more difficult to process than subject relative sentences, where the sentence elements have parallel roles in both clauses, such as “The child that bit the man admitted the error” (Just and Carpenter *ibid*). As Figure 5.1 illustrates, the average maximum pupil diameter for the more complex object relative sentence was higher than the subject relative sentence. It is also clear from Figure 5.1 that multiple data points are used in the trial-aggregated and fine-grained method, allowing the researcher to obtain detailed information on the timing of the pupil maximum, the onset of the change in pupil size, etc.

Just and Carpenter (*ibid*) use this form of measurement to complement their time-aggregated coarse method (see below). At first, it may appear desirable to follow their procedure and use both fine-grained and time-aggregated methods in our experiment. On the other hand, Klingner *et al.* (*ibid*) highlight that this form of fine-grained measurement can only deal with short, simple task components. While this was suited to Just and Carpenter’s (*ibid*) experiment, where tasks involved subjects reading aloud two short sentences that appeared on screen, it is not suited to the viewing of subtitled media, where subjects are likely to simultaneously deal with a wide variety of input from various semiotic channels along with each subtitle that appears. Moreover, the temporal proximity of some subtitles could result in the processing effort associated with a previous subtitle affecting the data reading for subsequent subtitles. For these reasons, the method applied in the present study was the time-aggregated coarse method, as explained in the following.

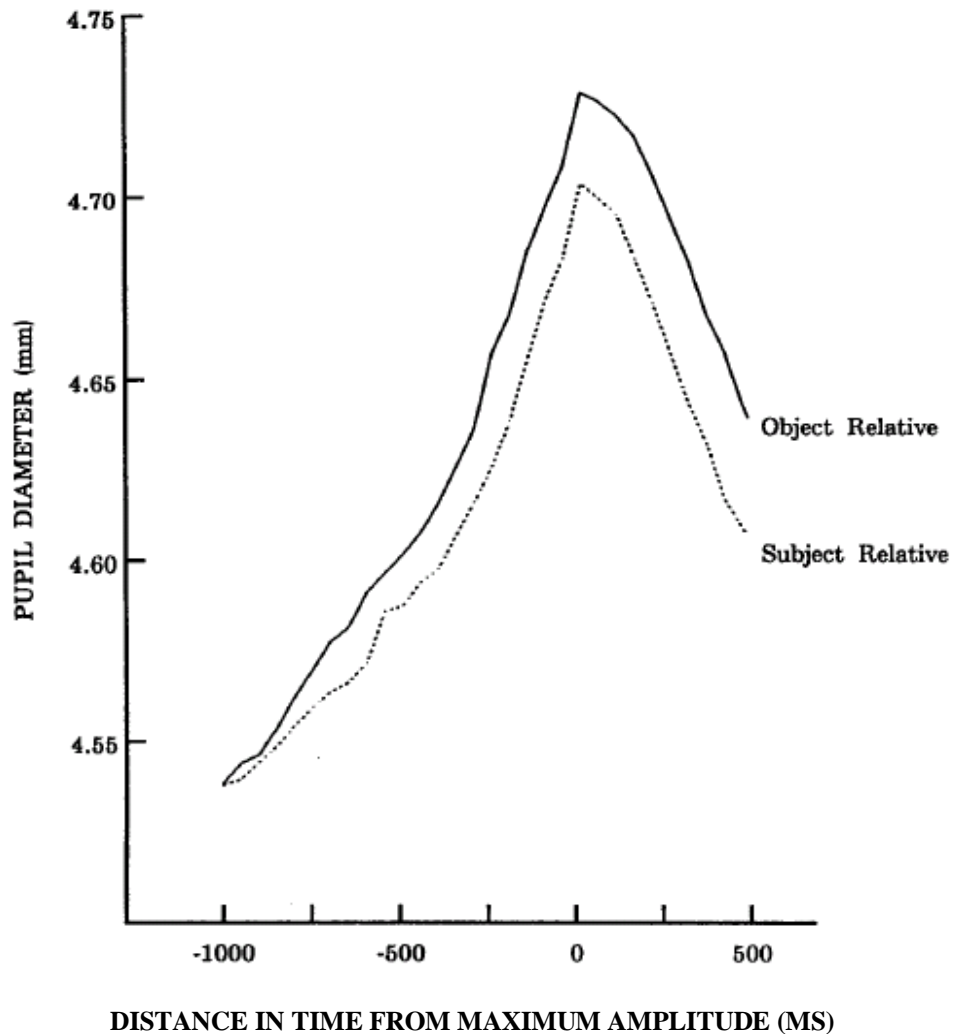


Figure 5.1. Waveform showing maximum pupil diameter calculated using trial-aggregated fine-grained method (Just and Carpenter 1993 p320)

Time-aggregated coarse method:

The time-aggregated coarse method used by Just and Carpenter (ibid) involved their averaging of the results for mean pupillary response obtained over 36 trials. This was measured by comparing a baseline pupil measure for each sentence with the higher measure that was expected to occur after the main clause verb in the sentence was encountered. The results were displayed in the format illustrated in Table 5.1. The results in the table reveal that the mean pupillary response was higher in the object relative sentences than subject relative ones, suggesting that more processing effort was required to process the object relative sentences, as expected.

<i>Sentence type</i>	Subject relative	Object relative
<i>Mean pupillary response</i>	.203 mm	.249 mm

Table 5.1. Mean pupillary response calculated using time-aggregated coarse method (adapted from Just and Carpenter 1993 p319)

When this form of presentation is compared with the waveform in Figure 5.1, a distinction between both methods becomes clear. The time-aggregated coarse method simplifies measurements to a unique figure (compared to a waveform with numerous data points), and ignores some of the complexities of the trial-aggregated fine-grained method, resulting in “a very robust but coarse measurement of cognitive load” (Klingner *et al.* 2008 p70).

A more refined measurement would be desirable for the instantaneous measurement of cognitive load that would be necessary for the interactive use of the eye tracker for adaptive subtitles (see Section 9.3). However, the aforementioned issues regarding the proximity of subtitles and the input from various semiotic channels rule out the use of the fine-grained method in this study. The purely diagnostic use of the eye tracker in the present study is suited to the time-aggregated coarse method, and is also more feasible in terms of time required to carry out the analysis. Furthermore, the use of mean figures allows the influence of various semiotic inputs on the pupillometric measurements to be reduced.

5.1.2.3 Unit of measurement used with pupillometric data in the present study

As there has yet to be any published study carried out using pupillometry with subtitled media, deciding on a unit of measurement for the pupillometric response was an important issue in the present study. All of the pupillometric studies mentioned above, with the exception of Schultheis (2004), Schultheis and Jameson (2004), and O’Brien (2008) use some form of baseline pupil measurement with which the changes in pupil size can be compared. Schultheis (*ibid* p57) explains the reasoning behind using a baseline as:

Über längere Zeiträume hinweg weitet und verengt sich die Pupille allein auf Grund allgemeiner physiologischer Vorgänge. So ist zu erwarten, dass der Pupillendurchmesser ein und derselben Person bei ein und derselben Tätigkeit an verschiedenen Tagen unterschiedlich groß ist. Solche Variationen treten auch in kürzeren Abschnitten als Tagen auf und es bestand die Möglichkeit, dass die beim Lesen der verschiedenen Texte erfassten Werte auch durch längerfristige Schwankungen mitbestimmt waren.

The pupil expands and dilates over longer durations of time due only to general physiological processes. It is therefore to be expected that the pupil diameter of the same person for the same task will be different on different days. This variation can also appear in shorter periods than days and there is a possibility that modulations caused over a longer period of time may contribute to the values obtained while reading different texts. [AT]

This has led to many studies using the measurement of percentage change in pupil diameter (PCPD) (e.g. Beatty 1982; Iqbal *et al.* 2005, O'Brien 2006). A baseline figure is generally used, and from this the change in pupil size while carrying out a certain task is calculated. This change is then calculated as a percentage of the baseline figure. Generally speaking, the higher the PCPD, the more cognitive effort is likely to have been required. To calculate the PCPD, O'Brien (*ibid*) did the following.

A baseline reading was calculated for each subject by taking a median pupil size measurement while subjects read a text. This baseline was then subtracted from the median left and right pupil size measurements taken from various segments of the translation task, each of which were categorised under one of four Translation Memory match types (no match, fuzzy match, MT match and exact match). The result of the subtraction was calculated as a percentage of the baseline, and then the results for left and right pupils were combined to provide a single PCPD measurement.

O'Brien (*ibid*) found in her study that in some instances the PCPD was a minus value, which would indicate that there was less effort required for the translation task than for the baseline task of reading, which goes against intuition. While PCPD was initially intended to be the unit of pupil measurement used in the present study, the occurrence of minus values in O'Brien (*ibid*) brought to light a possible drawback of using PCPD as an accurate unit of measurement. Furthermore, the variety of semiotic inputs and light intensities with which the subjects would be presented in the present study while viewing the subtitled AV

content were expected to make the establishment of an appropriate baseline task infeasible. In a more recent study, O'Brien (2008 p86) uses median pupil measurements instead of PCPD saying that she chose the median because she wanted to "know the exact midpoint of distributions and because individual differences could lead to some extreme scores that would distort the mean". In the present study, median pupil measurements for each subject are used for the same reason. The median provides a midpoint of the distribution of values for the period that each subtitle is on screen, which should lower the extent to which other semiotic inputs affect the pupillometric readings. These median values are then used to calculate means for statistical testing.

5.2 Eye tracking tool: *Tobii 1750* eye tracker

The *Tobii 1750* eye tracker was used to measure the eye movements and pupil size of subjects in the present study. Using corneal reflection technology, the eye tracker has near infrared diodes integrated into a 17-inch TFT¹⁴ monitor (see Figure 5.2). These diodes measure the location of the eye on the X and Y-axis of the screen as well as measuring pupil diameter. Firstly, the user's eye is illuminated to generate reflection patterns on their cornea. Then the reflection patterns are gathered using a video camera, along with the stance of the user. Digital image processing then extracts the pupil information from the video signal (Salojärvi *et al.* 2005).

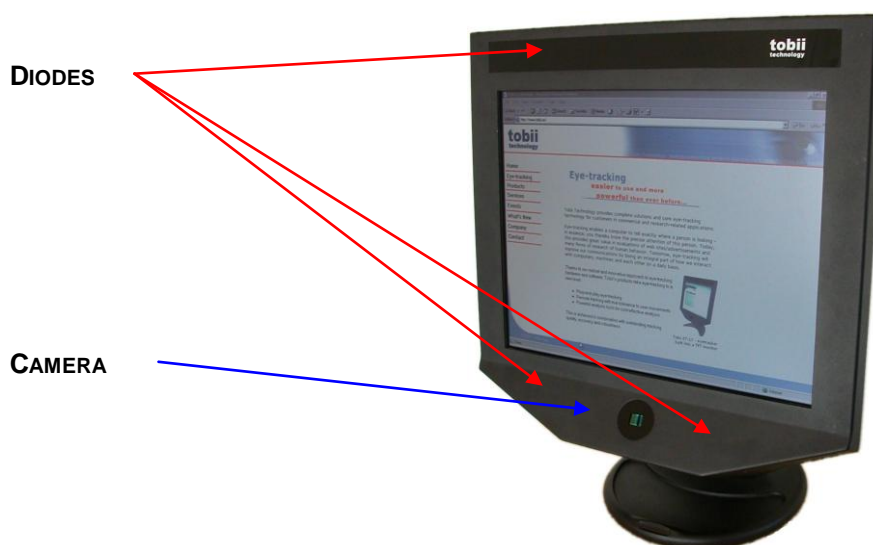


Figure 5.2. *Tobii 1750* eye tracker

¹⁴ Thin film transistor.

Eye trackers like this, which are integrated into monitors, allow for less intrusive eye tracking than head-mounted eye trackers, because the former allows subjects to look at the screen in a more normal manner.

5.3 Eye tracking software: *ClearView*

When preparing any stimuli for use with the eye tracker, the analytical software being used must be taken into account. The software used in this study was *ClearView*¹⁵ (Version 2.6.3), which has recording and analytical functions. It was the only piece of software used at the time of the experiment that interacted directly with the eye tracker. To prepare the excerpts for use with the eye tracker, the video was first converted to the .avi file format, and was then imported into *ClearView* as a multimedia stimulus. The .avi conversion was necessary because *ClearView* requires that video stimuli are in the .avi format, and DVD video is originally in the .vob file format.

5.3.1 Recording data using *ClearView*

Before showing subjects any stimuli, the eye tracker must be calibrated to their eyes. The calibration is a means of determining the likely quality of a pupil data recording for a specific subject. The *ClearView* calibration involves the subject focusing on and following a blue dot around the screen to nine data points. A graphic is then displayed showing the quality of the data gathered at each of the data points. As Figure 5.3 illustrates, a “good” calibration is indicated by few or no lines outside the small circles found at each of the nine data points, while a “bad” calibration is indicated by many lines leaving the circles. If no data is gathered for a certain point, the circle is surrounded by a red square.

¹⁵ *ClearView* has since been replaced by the updated software package *Tobii Studio*.

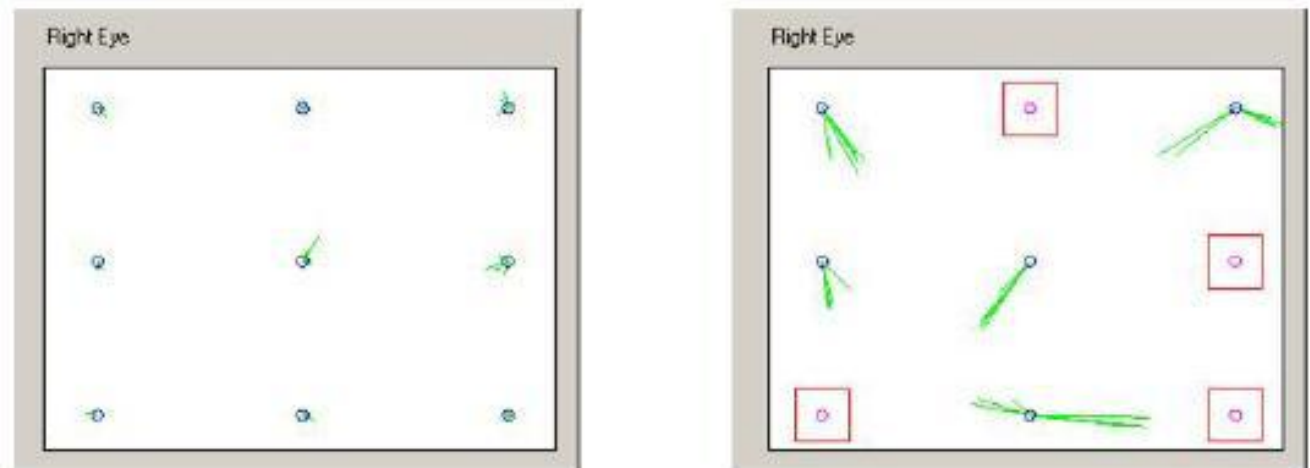


Figure 5.3. Example of "good" (left) and "bad" (right) pupil calibration in *ClearView* (Moran 2008)

Once the calibration has been completed, the subjects are shown the stimuli and a recording is made. After the recording has been made, data analysis can be carried out (see Appendix A for an outline of the process for preparing stimuli for analysis).

If using fixation-based measurements in *ClearView*, a fixation filter must be applied. The fixation filter defines the maximum distance between two points for them to be considered belonging to the same fixation and the minimum time for which the subject's gaze needs to be within the radius to be considered a fixation. For the current study the fixation was set to a radius of 30 pixels and a duration of 100ms, which is the recommended filter setting for mixed content (i.e. consisting of both text and images).

5.3.2 Data output

Data collected using *ClearView* can be exported in a number of ways and formats. As there is an extensive range of these formats, only those used in the experiment in this study will be outlined here: the Area of Interest (AOI) statistics template and hotspot diagrams.

5.3.2.1 Area Of Interest (AOI) statistics template

An *area of interest* (AOI), as the name suggests, is a user-specified area of the scene that is of interest to the given study. AOIs can be used in later analysis,

and fixation data may be gathered about them after they are defined in the program. AOI data can be exported to a prepared *Excel* template using *ClearView*. Figure 5.4 shows the typical layout of such a template. From this template, data can be displayed for multiple scenes, subjects and AOIs. The measures that can be calculated from this template are fixation count (the number of fixations in an AOI), gaze time (the total amount of time spent in an AOI in *ms*), average fixation duration (the gaze time divided by the fixation count in *ms*), and the time to first fixation (the amount of time it took to first fixate on an AOI in *ms*). In this study, fixation count, gaze time and average fixation duration measurements were used to calculate the dependent measurements for the fixation-based data analysis.

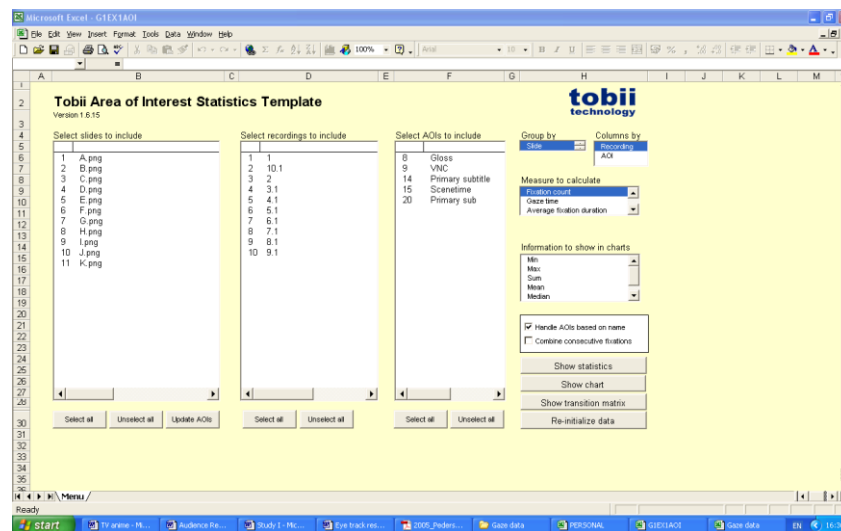


Figure 5.4. Screen capture from *ClearView* statistics template in *Excel*

These measurements were copied into a separate *Excel* spreadsheet, where the data for each subject was shown alongside the number of lines in the subtitle and whether there was a pop-up gloss present (indicated with a 'G' in the extras column in Table 5.2). The example in Table 5.2 below shows an extract from the spreadsheet for the measurement of fixation count, which was used to calculate the fixation probabilities for the subtitles.

Subject #	1	3	4	6	7	8	10	Lines	Extras
Subtitle									
A	3	5	5	4	3	5	2	1	
B	5	10	7	10	4	9	7	2	
C	9	19	12	16	13	13	12	2	
D	5	4	3	2	3	2	2	1	
E	12	12	9	11	11	9	9	2	
F	3	3	2	3	2	6	1	1	
G	5	8	8	7	7	7	5	2	G
H	5	11	11	5	13	12	10	2	G
I	3	2	1	1	2	2	2	1	
J	1	2	0	3	1	1	1	1	
K	0	3	5	4	3	2	3	2	

Table 5.2. Extract from *Excel* spreadsheet showing fixation count for selected subjects

5.3.2.2 Hotspot images

A hotspot image is a static representation from a scene with a heat map superimposed on it (see Figure 5.5). Hotspot images were exported to bitmap format (.bmp) files for each subtitle. The heat map can be set to show a graphical representation of several aspects of the eye behaviour, such as time spent in an area or average fixation duration in an area, but in this study the colours represent the number of fixations in an area.



Figure 5.5. Example of hotspot image, coloured areas representing areas that were fixated upon

Hotspot images may be exported for individual subjects, or groupings of subjects. The hotspot images used for each excerpt in the present study were combined for each subject who saw an excerpt. In the hotspot image, most fixations were recorded on the red coloured areas, less on the yellow areas, and the least on the green areas. No fixations at all were recorded for the areas that remain uncoloured. Hotspot images were used for qualitative analysis of eye fixation data, and while it was not feasible for reasons of practicality to carry out a full qualitative analysis on all hotspot images in the current experiment (totalling 404), they were useful as a means of explaining possible reasons for inconsistencies in the data collected by other means, e.g. questionnaire data.

5.3.2.3 Strengths and weaknesses of using *ClearView* with subtitled stimuli

While *ClearView* does have many capabilities as far as statistical data is concerned, it does not have the functionality for some of the calculations necessary in the current experiments. This meant that several fixation-based measurements had to be calculated separately by the researcher using *Excel*. The measurements that were calculated in this way were percentage of gaze time in subtitles, word fixation probability and percentage of skipped subtitles. An explanation of how these measurements were calculated is provided in Section 7.2.4.

Another disadvantage of using *ClearView* from the perspective of this study is that it does not provide any applications for dealing with pupil data, so all manipulations of pupil data had to be carried out manually, which was relatively time consuming when compared with fixation-based and questionnaire data, especially considering the vast volume of pupil size readings which were taken. O'Brien (2006) notes one of the difficulties in using pupillometric data when she highlights the difference in time-stamp formatting in the .avi video output from *ClearView* and in the *Excel* files. The video recordings of subjects eye-path can be analysed in *ClearView* in the minute:second:millisecond format, while raw data exported to *Excel* files is accompanied by a time-stamp given in milliseconds, necessitating a conversion. While *Tobii Studio* does have

increased functionality with fixation-based data, the pupillometric data is still only exportable in a plain text file or *Excel* worksheet file.

5.4 Summary

In this chapter the applicability of the eye tracker as a tool to measure processing effort was discussed, with an emphasis on the two types of measurement that are used in the present study, fixation-based measurements and pupillometric measurements. The four fixation-based measurements that are used were highlighted: percentage of skipped subtitles, percentage gaze time in subtitle area, mean fixation duration in the subtitle area and word fixation probability. Pupillometric measurements were discussed and the suitability of a time-aggregated coarse method for dealing with the pupil data in the present study was argued. The median pupil size was selected as the unit of measurement to be used with the pupillometric data, calculated for the period that each subtitle was on screen.

The chapter also provided a description of the eye tracking hardware and software used, describing some of the shortfalls of *ClearView*, and the format in which the data was exported from the software package. Before going on to discuss the methodological considerations and design of the experiment used in the present study in Chapter 7, Chapter 6 provides a critical overview of some of the relevant empirical research that has been carried out to date on the perception of translated AV content.

CHAPTER 6. **EMPIRICAL PERCEPTION** **STUDIES IN AVT**

Previous chapters have discussed the abusive procedure that is the focus of the present study, the pop-up gloss, and the source text used in the experiment described later in the thesis. The relevance theoretical notions that are used to analyse perception, PCEs and processing effort, were discussed, as well as the data collection methods used. This chapter presents an overview of the current state of affairs concerning empirical studies of perception in AVT.

Perception studies cover a wide range of study areas and they are truly interdisciplinary, with research interest coming from fields including literary studies (e.g. Richards 1929, Holub 1984), film and media studies (e.g. McQuail 1994, Austin 2005, Lay 2007, Santos 2007), communication studies (e.g. de Linde and Kay 1999a, 1999b) as well as TS (e.g. Fuentes Luque 2003, Gambier 2003, Chiaro 2004, Widler 2004). Studies of perception in TS fit into the fourth of Chesterman's (2005) four levels of translation research outlined as:

- i.) The cognitive level, which focuses on the decision-making processes in the translator's mind, quality of the translation task and the influence of the translator's cognitive environment on the translation process.
- ii.) The cultural level, which focuses on ideology, cultural identity, power relations and ethics in translation.
- iii.) The textual level, which is concerned with the text as an entity, relations between translations, their source texts and parallel non-translated texts in the target language.
- iv.) The sociological level is concerned with research on the translation market, the social status of the translator, translation as a social practice and recipients of the translation and their observable actions.

Falling under the sociological category of translation research, perception studies focus on the recipients of translations. In order to be considered an empirical study, a study must comply with one of the general aims of empirical research, which Chesterman (2000 pp10-11) describes as being:

to provide new data (such as a new corpus, or a detailed case study) on which hypotheses can be tested; to test an existing hypothesis (and perhaps refine it); to propose a new hypothesis (and justify it); to propose a new way of testing a hypothesis (or of generating one); to suggest new connections between hypotheses, i.e. a new theory.

The present study aims to provide new data on viewer perception of subtitled TV anime by employing a new method of testing hypotheses combining fixation-based and pupillometric data from the eye tracker with questionnaire data. As such it forms empirical research where the research design is informed by a critical review of prior empirical perception and reception research in AVT.

Despite the current popularity of AVT studies, there are still many areas that have remained unexplored (for a discussion of some of these see Gambier 2005). One area where the paucity of research has been highlighted (Antonini 2005, Gambier 2003, Gottlieb 2005b) is the reception of translated audiovisual products, although this too is beginning to change, with more research on the topic appearing, as the review below highlights.

Gambier (*ibid*) comments on the lack of empirical studies into reception in screen translation, despite the references that are often made to viewers, consumers and users. He notes that the paucity of research in this area could lead to a situation where:

...translators can only aim at a potential target audience whose profile they inevitably construct on the basis of their own stereotypes and prejudices; they further have to proceed on the assumption that their own individual social and cognitive environment is also that of the 'average' viewer (Gambier *ibid* p186).

The subtitler does not necessarily have to assume their social and cognitive environment is that of the average viewer, as they can change their translation strategy with an aim to suit viewers with different social and cognitive environments to their own. However, for the most part, Gambier's thought echoes Kovačič's (1995 p379) description of the "ideal viewer" from an author or literary critic's perspective, as one that will display:

...all the necessary cognitive, personal, social and other dispositions to receive a literary work exactly as it was conceived and intended by the author or as it is received by the critic.

Gambier (2003) asserts that the use of survey and experiments for investigating audience perception of translated texts is important to underline the linguistic and cultural responsibility of the subtitler. He describes both methods; with the survey method involving the use of questionnaire or interviews to elicit viewer responses to subtitled content, as well as experiments which “could provide insights into the effects of particular subtitle features...” (Gambier *ibid* p187).

6.1 Perception studies in AVT: Review

Although there are some studies into audience perception by translation scholars from the 1990s, the majority of these studies tended to focus on theoretical analyses, rather than empirical research, as Fuentes Luque (2003) illustrates. The number of empirical studies in the field of AVT has increased, and recently a cluster of research has emerged using eye trackers to empirically investigate the perception of translated AV content (Caffrey 2007, Caffrey 2008a, Caffrey 2008b, Caffrey 2008c, Moran 2008, Perego and Del Missier 2008).

Below is a review of relevant empirical studies carried out into the perception of translated AV content. Gottlieb (1995) uses two types of data collection for the study of viewer responses, consecutive and simultaneous, both of which are applicable in perception studies. Consecutive data collection methods include questionnaire and interview, while some simultaneous data collection methods that have been applied to date in AVT research include interactive tasks, observation and eye tracking. In the next section several relevant topics that have been researched in AVT will be reviewed according to the type of data collection they employed.

Consecutive data collection:

- Subtitle speed (Jensema 1998)
- Translated humour (Chiaro 2004, Antonini 2005, Bucaria and Chiaro 2007, Chiaro 2007)
- Culturally marked items (Bucaria 2005, Antonini 2007, Antonini and Chiaro 2005/2008)

Simultaneous data collection:

- Position of captions (Caffrey 2007)
- Subtitle speed and timing (Jensema *et al.* 2000a, 2000b)

Consecutive and simultaneous data collection:

- Errors in subtitling (Gottlieb 1995)
- Translated humour (Fuentes Luque 2003)
- Linguistic variation (Moran 2008)
- Segmentation quality (Perego and del Missier 2008)

6.2 Studies employing consecutive data collection

6.2.1 Questionnaire

Toury (1991) argues that questionnaires are an appropriate method for product-oriented empirical studies in translation, and most of the studies into the perception of translated AV content use of some form of questionnaire, with some using questionnaire alone as their tool for data collection.

6.2.1.1 Subtitle speed

Jensema's (1998) large-scale study was devised to determine viewer reactions to different subtitling speeds on TV (96, 110, 126, 140, 156, 170, 186, and 200 WPM¹⁶). The study used data from 578 subjects, consisting of deaf, hard of hearing and hearing people.

The above study consists of four parts: i.) a background questionnaire asking questions on topics including gender, degree of hearing loss, family status and

¹⁶ These speeds are all given in Words Per Minute, so a speed of 170 WPM means that on average 170 words will appear on screen per minute.

level of education, ii.) a simple vision test to ensure that the subjects were physically able to read the subtitles on screen, and iii.) two practice excerpts where the subjects were supervised as they categorised the videos according to the scale below¹⁷ (Jensema 1998 p320):

Too fast: Captions should be slower. Hard to read the captions. I miss some words.

Fast: Captions should be slightly slower. Captions should be on the screen a little longer.

OK: Caption speed is comfortable to me.

Slow: Captions should be slightly faster. Captions are on the screen a little too long.

Too slow: Captions should be much faster. I am bored while reading them.

Finally, subjects were shown 24 excerpts and asked to rate them on the same scale. The results showed that the mean subtitle speed that was rated as *OK* was about 145 WPM, which Jensema (ibid p324) notes is close to the 141 WPM mean subtitle speed he claims is found in TV programmes. The results also suggested that viewers were capable of adapting to increased subtitle speed, with significant difficulties only becoming apparent when the subtitle speed exceeded 170 WPM.

The scale used in the above study was adopted in the present study, although the explanations cited above were disregarded in the final version of the questionnaire for reasons explained in Section 7.6.3.3.

6.2.1.2 Translated humour

Chiaro (2004) describes the results of a study into the perception of translated verbally expressed humour (VEH). The study was carried out using an online self-reporting questionnaire, and was administered at the beginning of 2004 to a random sample of Italians and received 65 valid responses (relevant to subjects' perception of VEH). The subjects were randomly shown a selection of excerpts from episodes recorded from several well-known North American TV series dubbed in Italian that were broadcast during a three month period in 2002 from all Italian terrestrial channels (Chiaro ibid p35). The excerpts were accompanied by a brief description of the series and current situation for

¹⁷ Note that Jensema uses the term captions to refer to subtitles, as is common practice in North America.

contextualisation. After watching the excerpt, subjects were asked to write what they understood of the VEH in a word pad.

Chiaro (2004) carries out a qualitative analysis of the subjects' answers and argues that in some cases poor translations cannot be entirely blamed for instances where VEH is misunderstood, instead highlighting that when "humour depends on the recipient's encyclopaedic knowledge, sometimes it is successfully recognised as such, and sometimes it is not" (Chiaro *ibid* p49).

In a study of Italian viewers' perception of a subtitled episode of the English TV series *Father Ted*, Antonini (2005) tries to determine whether the socio-cultural differences in relation to Catholicism in Italy would trigger annoyance at humour aimed toward the church, and whether the audience would be more amused by visual humour elements than puns and punch lines mainly based on verbal elements. She carries out this study with the use of a questionnaire based on the psychologist Ruch's (1992) 3WD¹⁸ test on humour appreciation.

Subjects were shown an episode of *Father Ted*, and were then asked a general question on subtitling. Afterwards, they watched nine excerpts that contained examples of visual and verbal humour, and humour aimed at the Catholic Church. Subsequently, they were asked to rate the excerpts on two six-point Likert scales aimed at judging how funny and how aversive they found each of the excerpts. They were also asked whether they understood the semantic content in question for each excerpt before they were asked to explain it briefly. The study found an explicit correlation between low ratings of funniness and partial understanding.

The results of the study reveal that viewers had difficulties in understanding the verbal humour through the subtitles, and found it less funny than visual humour. It was also found that almost half of the sample subjects recreated puns, despite the fact that they were not translated in the subtitles. One possible

¹⁸ The 3WD, or 3 Witz-Dimensionen test, is a humour test used to classify respondents' personality traits according to their responses and reported preferences to three types of humour: incongruity, nonsense and sex.

reason suggested for the recreation of puns was the presence of canned laughter in the excerpts, indicating that something funny was happening.

The study by Bucaria and Chiaro (2007) focuses on three features of dubbed programmes, culture specific references, language specific features and what they term “borderline features”, or those aspects that cross both language and culture, such as VEH, songs, gestures and allusions.

A total of 150 questionnaires were distributed to Italians, in three different groups: i.) self-selected respondents chosen by the researchers through personal networking from the area of Cesena and Forlì, ii.) respondents who work in the field of dubbing, in Milan and Rome iii.) experts in the area of linguistics, translation and media studies, mainly from the areas of Bologna, Reggio Emilia and Genoa. While surveys were administered face-to-face wherever possible, a “considerable” number were sent and collected by post. Subjects were shown a tape containing twenty excerpts from a selection of TV programmes dubbed into Italian and asked to answer questions on them. A variety of questions were used, with preference for multiple choice and yes/no questions to increase the ease of classification, although some open questions and metric graphic rating scales¹⁹ were used. 87 questionnaires were returned, the results of which once again revealed that subjects’ understanding of the content suffered when there were gaps in the viewers’ encyclopaedic knowledge.

Chiaro (2007) uses questionnaire to compare the humour response of 22 British and 34 Italian subjects to excerpts from several popular English-language comedies as well as one excerpt from an Italian-language comedy. British subjects saw the excerpts in the original version, while the Italian subjects saw them dubbed or subtitled in Italian. Some excerpts showed examples of visual humour, while others used VEH. Chiaro tests the hypothesis that “members of Culture A exposed to VEH which has been translated into Language A (Italian) from Language B (English) will display a less positive reaction to the same VEH

¹⁹ These are scales where a 10 cm line is placed on the questionnaire, with two polar opinions at each end (e.g. “not at all” and “completely”), and the subject marks the location where their opinion lies on the line.

than members of Culture B (UK) exposed to the same VEH in Language B (English)” (Chiaro 2007 p144).

A bilingual questionnaire was used, which was divided into two sections. The first section asked subjects for socio-demographic information, as well as a self-assessment of their sense of humour and their mood at the time of filling out the questionnaire. The second section asked questions about the excerpts. Subjects were asked to rate the funniness of each excerpt on a seven-point scale ranging from “it irritated me” to “it made me split my sides laughing” (Chiaro *ibid* p143). In excerpts where VEH was used, subjects were asked to rate their appreciation of the VEH on a metric graphic rating scale from one to ten.

The results of the statistical tests carried out on the data revealed that British subjects found all of the excerpts funnier than the Italian ones did, with the exception of the Italian comedy, which the British found less funny. The research hypothesis was confirmed with more than 60% of the examples of VEH resulting in a significantly higher level of appreciation from British subjects.

6.2.1.3 Culturally marked items

Bucaria (2005) used a questionnaire to investigate the perception of nonverbal items in AV content dubbed into Italian. Subjects watched ten excerpts from various sources, each of which contained a “culturally specific nonverbal element”. They were asked to rank their understanding of the nonverbal item on a scale of one to ten before answering an open question on the item, which was subsequently coded as a yes/no variable. Similar to Antonini’s (*ibid*) study, the results from this study revealed that declared understanding (79%) of the nonverbal items was more than double the actual understanding (31%) of the items. Bucaria (*ibid*) concluded her study by noting the role translation plays in guiding viewers to recognise nonverbal cultural references in AV content, and asking “to what extent can audiovisual texts be manipulated to facilitate the understanding of non-verbal, culture-specific elements?”

Bucaria’s (*ibid*) study is similar to the present study in that it also focuses on the perception of CVNCs. However, she uses only one subtitle set with each

excerpt, while the present study uses two subtitle sets. The use of two subtitle sets was necessary in the current study in order to examine the difference between the PCEs experienced by subjects who saw the same excerpts with or without pop-up gloss.

Antonini (2007) used a large-scale online questionnaire with 253 respondents to measure Italian audiences' perception of culturally marked items in dubbing. The questionnaire was advertised on a popular Italian website²⁰ between February and May 2004. As such the sample was drawn from the on average 12 million web users who access the website on a monthly basis (Antonini *ibid* p158).

There were 60 excerpts relating to what Antonini terms “culture-specific references”, all chosen from the same corpus used by Chiaro (2004). Each excerpt was introduced with a short synopsis of the episode it came from. After watching the excerpt (selected at random from the 60 available), perception of culture-specific references was measured by asking respondents to rate their understanding of the content on a graphic rating scale from zero to ten, and to explain what they had understood.

Again, the results of this study showed that more respondents declared understanding of the culture-specific references than actually understood them, indicating that “although Italian TV audiences believe and declare that they have understood most of these references, in reality they have not” (Antonini *ibid* p161). She suggests that developments in the multimedia sector may help to narrow the gap, with the possibility of including “the kind of added explanatory information that may help them understand and better appreciate what is happening on the screen” (Antonini *ibid* p165).

6.2.2 The questionnaire as a form of data collection

All of the above studies effectively use questionnaire as a means of measuring PCEs relating to various aspects of the AV content. The widespread use of questionnaire to measure comprehension of AV content make it an obvious choice to use in the present study. It was also the most practical way of

²⁰ <http://www.viriglio.it>

determining whether subjects understood the CVNCs and retained the subtitle content in the context of the current study. While interview could have been used, questionnaires provide more concise data, allowing for a more efficient collation of answers. The importance of efficiency when dealing with the questionnaire data becomes clear when the volume of data generated by the eye tracker is taken into consideration. The most prominent drawback of using consecutive questionnaires to measure PCEs is the possibility of subjects' memories distorting the results. So while a subject viewing an excerpt may realise that a character with a nosebleed is aroused, by the time the same subject answers a question about the nosebleed in a questionnaire they may have forgotten its meaning.

Questionnaires can also be used effectively to measure the processing effort experienced by subjects. However, the results of questionnaire-based scales of processing effort are inevitably biased by subjects' subjective opinions. Although the questionnaire in the present study did not directly ask about processing effort experienced²¹ in an attempt to prevent subjects consciously thinking about the effort they exerted while viewing the excerpts, it was considered necessary to include complementary measurements of processing effort. One reason for this is that the link between perceived speed of subtitles and processing effort has not been proved. Also, a series of short and simple subtitles with simple content may be perceived as being quicker than a series of long, complex subtitles that are shown on screen for longer periods of time. However, they may be less difficult to process.

For this reason, simultaneous measurements gathered with an eye tracker were deemed the most suitable complement to the consecutive measurement of processing effort provided by questionnaire because they provide an instantaneous, objective source of data, and protect against some of the negative effects of consecutive questionnaires. For example, subjects forgetting the subtitle speed (and providing a nonsense answer) or being self-conscious about their reading speeds (and consciously providing a false

²¹ Instead subjects were asked to rate the perceived subtitle speed, from which processing effort was inferred.

answer) would be likely to affect the answers to questions on subtitle speeds, thus corrupting the data.

The questionnaire types used in the above studies were adapted to measure the perceived and actual PCEs experienced by subjects in the present study. The questionnaire used in the experiment borrows heavily from Antonini's (2007) study²², asking subjects to rate their understanding of the CVNC before they describe the meaning of the CVNC, allowing for a comparison to be drawn between the perceived and actual PCEs experienced by viewers. The nature of the present study, where eye tracking was used in conjunction with questionnaire, meant that the sample size was smaller than the above studies, as each subject was tested individually in the lab where the eye tracker is housed, making the use of online and postal questionnaires infeasible. More details about the design and content of the questionnaire are discussed in Chapter 7.

6.3 Studies employing only simultaneous data collection

6.3.1 Eye tracker only

6.3.1.1 Positioning of subtitles

A small preliminary study was carried out on four undergraduate multimedia students as an opportunity for the researcher to become familiar with the eye tracker (Caffrey 2007). The study was designed to test the effects of the use of three types of non-dialogue subtitles on the reading time of dialogue subtitles in six excerpts from the commercial DVD releases of two TV anime series, *Doki Doki School Hours* and *あずまんが大王* (*azumanga daiō*, Azumanga Daioh 2004). Reading time was calculated by dividing the duration of time in seconds that the subtitle was fixated on by the number of words in the subtitle, resulting in the unit of seconds per word.

Caffrey (ibid) defined dialogue subtitles as subtitles translating the primary

²² This study is based on a previous experiment carried out by Antonini and Chiaro (2005), and published in Antonin and Chiaro (2009).

narrative dialogue, while non-dialogue subtitles were those that translated any other aspect of the AV content except the primary narrative dialogue, including secondary dialogue (e.g. type of subtitles used if two speakers are speaking at the same time, see *dual speaker* below), on-screen text and songs. Three positioning strategies used for non-dialogue subtitles were investigated in the study:

1. *captioning*, where the non-dialogue subtitle is found on top of the screen (see Figure 3.2)
2. *dual speaker*, where the non-dialogue subtitle translates secondary, background dialogue, found above the dialogue subtitle (see Figure 6.1)
3. *scenetime*, where the non-dialogue subtitle is placed close to the text it is translating (see Figure 3.3, bottom).



Figure 6.1. Screen capture showing a "dual speaker" subtitle (surrounded by red box) and a normal dialogue subtitle (*Azumanga Daioh* 2004)

The analysis of the eye tracking data suggested that the presence of non-dialogue subtitles decreases the reading time of subtitles, that is, the amount of time spent in the subtitle area. This difference was most pronounced when scenetime was used. This study provided further impetus to investigate the effects that pop-up gloss would have on the processing of subtitles by the viewer. However, whereas the above study used the measurement of seconds per word, the current study uses word fixation probability and mean fixation

duration on the basis of d'Ydewalle and de Bruycker (2007), as the latter two take more account of the fact that not all words are fixated upon, thus providing a more accurate value.

6.3.1.2 Subtitle speed

Jensema *et al.* (2000a, 2000b) carried out two studies using eye tracking. The first (Jensema *et al.* 2000a) used six subjects, three of whom were deaf and three hearing. It aimed to determine: i.) whether the presence of subtitles changed the way subjects viewed television, ii.) if viewing strategies differed among subjects, iii.) whether prior knowledge of the video content changed viewing strategy and iv.) how the subtitle speed influenced viewing.

Subjects were shown eight excerpts, and their eye movements were recorded. Several variables, such as subtitle speed and presence of subtitle, were used to make a qualitative analysis of the eye movement patterns.

The results suggested that people who view excerpts with similar content have similar eye movements, while the introduction of subtitles in the sequence drastically changes eye movement patterns, with viewers looking first at the middle area of the image, before shifting their focus to the start of the subtitle. They also found that upon the second viewing of the same excerpt, viewing strategies became even more similar. Furthermore, as subtitle speed increased, more time was spent in the subtitle area, possibly because of the prioritisation given to reading subtitles (Jensema *et al.* 2000a).

In an expanded study, Jensema *et al.* (2000b) recorded the eye movements of deaf subjects to determine the amount of time spent viewing subtitles. They focused on four factors, namely: i.) the percentage of time deaf people spent looking at subtitles; ii.) the influence of age, sex and education on the percentage of time spent looking at subtitles; iii.) the influence of programme type and subject on the percentage of time spent looking at subtitles; and iv.) the influence of subtitle speed on the percentage of time spent looking at subtitles. 23 deaf subjects were shown four 2.5 minutes long excerpts, which were divided into five segments, each subtitled at one of the following speeds: 100, 120, 140, 160, and 180 WPM.

The results revealed that on average, subjects spent 84% of the excerpt duration looking at the subtitle area, 14% looking at the rest of the image, with the remaining 2% off the video. They found that age and caption viewing time were not closely linked in their study, nor was there a significant difference between male and female subjects, which contradicts the results of d'Ydewalle *et al.* (1987), who found that female subjects spent significantly more time in the subtitle area than their male counterparts. Level of education was seen to have a significant correlation with the amount of time spent in the subtitle, although it was not clear from the data what these results meant in practical terms, resulting in a call for further research investigating “the relationship between education and percentage of time spent reading captions” (Jensema *et al.* 2000b p467). There was also little difference noted in the amount of time spent in the caption area between programme types or subtitle speeds. The most distinctive characteristic of their data was the differences between subjects regarding the amount of time spent in the subtitle area, suggesting that factors such as interest in the topic, prior knowledge of the subject and legibility of the subtitles may have had an effect.

6.3.2 Using eye tracker only as a method of data collection

All of the studies above use an eye tracker as their only tool for collecting data on subjects' perception of subtitled AV content. An eye tracker integrated into a desktop monitor is a helpful tool for generating data on viewer perception. Integrated eye trackers are unobtrusive, resulting in lower interference with viewer behaviour, and they also generate a large volume of objective data.

The gaze location data generated by the eye tracker can provide an objective insight into behavioural features of the perception of subtitled AV content such as subtitle reading speed, attentional distribution between subtitle and image, the order in which pictorial items are perceived and how often subtitles are refixated upon. However, in the context of the present study, the use of an eye tracker alone was not considered as a suitable form of data collection.

Whereas the eye tracker alone could conceivably be used to measure processing effort, the aim of the present study to measure processing effort *and*

PCEs meant it was necessary to use another data collection method to gather data on PCEs.

6.4 Studies employing simultaneous and consecutive data collection

6.4.1 Interactive task and questionnaire

Gottlieb (1995) reports on an experiment carried out in April 1995 investigating reactions to errors in subtitles. The responses of 123 deaf and hard of hearing subjects to 12 intralingually subtitled excerpts from Danish language television were recorded using questionnaire and a protest button.

The protest button was used to measure reactions to individual subtitles, while the questionnaire measured reactions to the excerpts as a whole. The questionnaire consisted of roughly half multiple choice questions while the rest were open questions or yes/no questions, based on the content from various semiotic channels in the excerpts. These included questions about the languages spoken in the excerpt, and about the content of the image and dialogue.

Preliminary analysis of the data suggested that the audience only reacted negatively to subtitles where dialogue misrepresentation was severe, such as the heavy cutting or distortion of information. This led Gottlieb (ibid) to the conclusion that even if only a few subjects note something wrong with a subtitle it may warrant further investigation. Faulty subtitles generally were not reacted to, and over half of the negative responses were to standard practices such as speech condensation. Subjects, even those who had full hearing loss and were therefore totally dependent on the subtitles for the transfer of information in the audio verbal and nonverbal channel, took in nonverbal visual information before subtitles. Questionnaire data revealed that subjects retained more nonverbal visual information than information contained in the subtitles. This finding adds further weight to the argument of the importance of the visual nonverbal aspects of AV content.

6.4.2 The “protest button” as a form of data collection

The protest button used in Gottlieb’s (ibid) experiment was attached to the subject’s chair and a computer recorded each button press. Subjects were asked to press the button when they found something in the subtitles confusing or erroneous. The collection of data using a protest button can be categorised as a task-based performance measure, where the subject must actively input data according to parameters set in the experiment.

At first glance, the use of a protest button in the present study may appear to be desirable, as it provides a simultaneous measure that could easily be linked to individual subtitles. It is also easy to explain to subjects, and the data generated is relatively easy to work with. For example, in the context of the current study, subjects could be asked to push the button when they feel that they require too much effort to watch part of an excerpt, and the number of button pushes depending on the presence or lack of pop-up gloss could be compared. However, Gottlieb’s (1995) experience with this form of data collection led him to believe that while some subjects did indeed react in the places they felt led astray, others tended to either produce a large output of button presses, or to slip into a non-responsive state.

This issue of either over- or under-responsiveness was considered to be one of the main reasons not to use such a method of data collection in the present study. The use of pop-up gloss was expected to cause difficulties for some (if not all) subjects at some stage while reading the subtitles, and similar reactions to those found in Gottlieb’s (ibid) results could be expected.

Task-based data collection also brings up some issues related to experiment validity when it is used in subtitle perception studies. Firstly, the authenticity of the study setting is affected, with subjects dividing their attention between the subtitled excerpts and the task of pressing the button. As well as this, if the viewer is being asked to react to a certain aspect of the audiovisual text, a disproportionate amount of attention may be paid to that aspect, once again resulting in a more artificial viewing condition.

Finally, the use of the protest button may interfere with readings obtained from other measures of cognitive processes, leading to confounding data. This can occur when the extra cognitive effort required to press the protest button influences the cognitive measurements related to the process under investigation (in the present case, watching subtitled TV anime). This issue is particularly relevant to the experiment in the present study, where pupillometry is used to measure processing effort.

6.4.3 Observation and questionnaire

Fuentes Luque's (2003) empirical study into the reception of translated humour uses three versions of a ten-minute segment from the Marx Brother's film *Duck Soup* (1933): the original English version, the dubbed Spanish version, and the subtitled Spanish version with original English soundtrack. One of each version was shown to separate groups, the original English version to a group of native English speakers, and the other two versions to native Spanish speakers. It was ensured that subjects in the Spanish-speaking group who saw the subtitled version did not speak any English, in an attempt to avoid "linguistic 'pollution' from the background original soundtrack in English" (Fuentes Luque *ibid* p296).

Similar to Chiaro's (2007) findings, the study suggested that viewer reaction to humorous elements in both Spanish versions was significantly lower compared to the level of positive reception of the original English version, which Fuentes Luque associates with the translation strategies used in the dubbed and subtitled versions. In conclusion, he agrees with Delabastita (1996 pp134-135, quoted in Fuentes Luque *ibid* p305) when he asserts: "the only way to be faithful to the original text (i.e. to its verbal playfulness) is paradoxically to be unfaithful to it". This could be construed as further support for the use of abusive subtitling, where traditional norms can be sacrificed in order to be "faithful to the original text".

6.4.4 Observation as a form of data collection

In his study, Fuentes Luque's (*ibid*) consecutive data was collected using a questionnaire and brief interview, while simultaneous data was generated by observing subjects as they watched the content under investigation. Subjects

were observed out of their range of vision and their reactions to the content were coded as one of the following: no reaction, smile, laughter or puzzlement.

Data collection by observation could be considered a suitable method to be used for measuring processing effort in conjunction with eye tracking. Firstly, like eye tracking data, observational data should be collected on an individual, rather than a group basis, in order to avoid what Fuentes Luque (2003 p296) refers to as “‘solicited’ or ‘suggested’ reactions” from other subjects. Secondly, some eye trackers such as the *Tobii 1750* have a function that allows for a video recording of the viewer to be taken. And thirdly, data collected by observation should have little or no influence on other cognitive load measurements, because of the relative unobtrusiveness of this form of data collection.

However, there are several drawbacks of this method of data collection which led to the decision not to adopt it in the experiment used in the present study. The first, and most striking, is the difficulty of gaining adequate insight into the “black box” of subjects’ minds by simply observing their external reactions. While the ostensive display of a reaction is more likely in humour studies, the focus of this study makes the viewer response less easily observable and gradable. The use of observational scales, such as the one used by Fuentes Luque (*ibid*) to measure varying degrees of humour²³, also has drawbacks. Firstly, the observation method can run the risk of the researcher’s subjective judgement influencing the results. Secondly, the reactions of viewers in many situations will not be scalable (e.g. not everyone laughs aloud when they find something funny, thus reducing the validity of between-subject comparisons made with the data).

In the context of the current study, one clear shortcoming is the lack of any widely accepted comparable scale of facial expressions that can be used to measure the degree of processing effort a viewer is experiencing. The lack of any adequately quantifiable, externally visible measure of processing effort means that the applicability of this data collection method in this study is too

²³ Fuentes Luque (*ibid*) categorised all reactions as a.) no reaction, b.) smile, c.) laughter, or d.) puzzlement.

limited to warrant its use. The eye tracker on the other hand provides a more objective and independent means of data collection, and is less easily affected by possible bias of the researcher.

6.4.5 Eye tracking and questionnaire

6.4.5.1 Linguistic variation

Moran's (2008) study aims to determine whether linguistic choices in subtitling facilitate the reception of subtitled material. She carried out an experiment with 16 subjects who viewed an excerpt with two different subtitle sets. Using a *Tobii x50* eye tracker, two variables were measured for their effects on subtitle reading: word frequency and use of cohesive ties. The measurements used from the eye tracker were average fixation duration and number of fixations.

For the first variable, word frequency, subjects were shown the same excerpt in two conditions, one with a word more frequently used in the movie, and the other with a less frequently used word. She provides the example in Table 6.1 below to illustrate the difference between the high frequency and low frequency conditions.

<i>Condition</i>	<i>Subtitle</i>
High frequency	As a result he thinks she has a heart defect .
Low frequency	As a result he thinks she has a heart problem .

Table 6.1. Same subtitle in high frequency and low frequency conditions (Moran 2008)

The results showed that average fixation duration on subtitles with the high frequency words was lower than in the low frequency condition; and that the number of fixations on the image was higher in the high frequency condition. This leads her to suggest that the use of high-frequency words could allow the viewer to allocate more attention to the image.

The second variable, cohesive ties, was dealt with as follows. Again subjects were shown the same excerpt twice. In one condition, cohesive ties were added, such as conjunctions, substitution, repetition and ellipsis, and in the other they were not, as is illustrated by the example in Table 6.2.

<i>Condition</i>	<i>Subtitle</i>
Low-cohesion	Today, Bretodeau won't buy anything.
High-cohesion	But today, Bretodeau won't buy a chicken .

Table 6.2. Same subtitle in low-cohesion and high-cohesion conditions (Moran 2008)

In this example, the conjunction “but” and the repetition of the word “chicken” from an earlier subtitle were used. The results from this section revealed that mean fixation duration in the low-cohesion version was significantly higher than it was in the high-cohesion version. However, there was no significant difference found in the number of fixations in the image.

The methodology applied in Moran’s (2008) study differed from that of the present study in that the former used the same subjects who were shown the same excerpt in both experiment conditions. While that did allow for the controlling of all the semiotic elements in each excerpt, as well as for individual differences in viewing patterns, this was avoided in our study. The problem with the same subject viewing the same excerpt twice becomes clear when we consider that the subject’s viewing is likely to be different the second time around, which was highlighted by the results of Jensema *et al.*’s (2000a) study. Jensema *et al.* (ibid) found that when subjects were shown the same excerpts a few days after the first viewing they had an idea what to expect in the content and eye movement patterns standardised across subjects. In order to avoid this contamination of data influencing the internal validity of the experiment, it was decided that no subject would see any excerpt more than once in the present study.

6.4.5.2 Segmentation quality

Perego and Del Missier (2008) presented the first part of results of their study into the effects of line segmentation and subtitle presentation on the comprehension of film content, picture recollection and vocabulary retrieval in an attempt to determine how existing subtitle norms meet the needs of viewers. They prepared an experiment with two conditions, one where “high-quality” line segmentation was used in the subtitles, and one where “low-quality” segmentation was used. 41 native Italian-speaker volunteers watched a 15-minute excerpt from a Hungarian film subtitled in Italian. Each excerpt contained half of the subtitles with high-quality segmentation and half with low-

quality segmentation. Some subjects saw the first half with high-quality segmentation and the second with low-quality, while the rest saw it in the reverse order.

Questionnaires were used to test readability on a scale of one to seven, to measure comprehension and to measure the recollection of images from the excerpt. While the eye tracking aspect of the results has not been published yet, the questionnaire results suggested that there was no significant difference between high quality and low quality segmentation in either readability scales, comprehension or image recognition. Instead, the results suggested that individuals' attentional control could have more of an effect on task performance than segmentation quality.

6.4.6 Eye tracker and questionnaire as a method of data collection

The survey of methods used in prior relevant studies suggests that the combination of data collected using an eye tracker and questionnaire provide a robust method of analysing the perception of subtitled AV content. The combination of these data collection methods makes it possible to carry out an effective analysis of both aspects of perception under investigation in the present study.

The use of a questionnaire provides an efficient means of measuring PCEs, generating data that can be relatively easily collated, which is an aid in the process of statistical analysis. Questionnaire can also be used to measure processing effort, albeit with some significant shortcomings, including the influence that subjective opinions of subjects may have on their answers. In order to reduce the subjectivity of the questionnaire data in terms of processing effort, the complementary use of data gathered from subjects simultaneously with an eye tracker while they viewed the excerpts was considered to facilitate a more holistic analysis of processing effort. Two forms of eye tracking data are used, fixation-based measurements and pupillometric measurements, each measurement acting as a means to validate the other as an effective measurement of processing effort. The measurement validity was particularly

relevant for pupillometric data, which has yet to be used as a measure of processing effort with subtitled AV content.

The combination of these three data types (questionnaire, fixation-based and pupillometric measurements) has not yet been used in a study of the perception of subtitled AV content. Thus it was necessary to give thorough consideration to the methodological issues concerning the present study, and the design of the experiment that would be carried out. The first of these issues that will be discussed is the sample size used in the studies reviewed above, while further issues will be discussed in the following chapter.

6.5 Sample size in reviewed AVT perception studies

The data in Table 6.3 shows the number of subjects used in the related studies reviewed above, illustrating the variation between studies, ranging from four (Caffrey 2007) to 578 (Jensema 1998). There appear to be several factors that were related to sample size. The first of these is the distinction between academic and industry or government sponsored research, with two of the largest scale studies being carried out in cooperation with either a television station (Gottlieb 1995) or government bodies (Jensema *ibid*). Excluding the three exceptionally large-scale studies just mentioned (Gottlieb *ibid*, Jensema *ibid*, and Antonini 2007), the mean number of subjects in the studies reviewed above is 36, although there is quite a large variation in the numbers between studies, with a standard deviation of 25.32.

Another factor that appears to be related to the number of subjects is the data collection method that was employed. The use of online questionnaires distributed over the internet resulted in a high response rate to Antonini's (*ibid*) study. On the other hand, studies that employed an eye tracker (see area on Table 6.3 shaded in dark grey) tended to use fewer subjects than the other forms of data collection, with median values of 16 and 65 subjects respectively. This is likely due to the large volumes of data generated by the eye tracker, and the implications this has on the time necessary to analyse the data. The number of subjects recruited for the present study was set at 20, as it was close to the median number of subjects used in the reviewed studies using questionnaire and eye tracker. While there have been some larger-scale

studies using over 100 subjects carried out using an eye tracker by commercial firms, the use of such a scale of subjects is unviable in academic research without substantial financial backing.

Data collection method	Focus of study/author	No. of subjects/ respondents
Questionnaire	<i>Subtitle speed</i> Jensema 1998	578
	<i>Translated humour</i> Chiaro 2004 Antonini 2005 Bucaria and Chiaro 2007 Chiaro 2007	65 32 87 56
	<i>Nonverbal items</i> Bucaria 2005 Antonini and Chiaro 2005 Antonini 2007 Antonini and Chiaro 2009	36 253 ²⁴
	Task and questionnaire	<i>Errors in subtitles</i> Gottlieb 1995
Observation and questionnaire	<i>Translated humour</i> Fuentes Luque 2003	30
Eye tracking	<i>Positioning of subtitles</i> Caffrey 2007	4
	<i>Subtitle speed and timing</i> Jensema et al. 2000a Jensema et al. 2000b	6 23
Eye tracking and questionnaire	<i>Linguistic variation</i> Moran 2008	16
	<i>Segmentation quality</i> ²⁵ Perego and del Missier 2008	41

Table 6.3: Number of subjects in reviewed studies organised according to data collection method and topic of research

6.6 Summary

This chapter provided an overview of the key relevant research that has been carried out to date into the perception of translated AV products. The recent surge in interest in using the eye tracker as a data collection tool was highlighted, and the lack of any research using pupillometry with subtitled AV content was noted. The topics and methodologies from the studies reviewed here provide a solid base from which to develop the methodology for the

²⁴ Antonini and Chiaro (2005), Antonini (2007), and Antonini and Chiaro (2009) were based on the same experiment. For this reason only one figure is given.

²⁵ Perego and del Missier's eye tracker data was not yet analysed at the time of the presentation at the conference.

present study which focuses on the little explored area of the perception of translated anime in general, not to mention the effects of pop-up gloss on the same. The next chapter focuses on the methodological considerations that informed the design of the experiment in the present study.

CHAPTER 7. **METHODOLOGICAL** **CONSIDERATIONS AND EXPERIMENT** **DESIGN**

Chapter 5 discussed the use of the eye tracker as a tool for measuring processing effort, and Chapter 6 reviewed relevant prior empirical research that has been carried out into the perception of subtitled AV content using data collected with an eye tracker. Given the novel application of the eye tracker, the present study considered the use of a pilot study as a vital step to prepare for the design of the main experiment. This chapter begins with a brief description of the pilot study, followed by the methodological considerations and the experiment design used in the present study.

7.1 Pilot study

The pilot study was carried out in DCU in March 2007 as a formative experiment with the primary goal of increasing familiarity with the eye tracker and *ClearView* software, in addition to ascertaining the validity of combining data collected with an eye tracker and questionnaire. It also provided an opportunity to develop and test a design template for the main experiment. A formal ethical approval step was not pursued in this study, mainly due to the fact that the internal DCU survey using a student portal web site necessitated approval from the Student Union President of DCU to ensure that students' rights were not infringed upon. This was considered enough approval for the pilot study, although a more strict ethical approval process was followed for the main experiment (see Section 7.6.2).

7.1.1 Aim

In addition to the above motivation, the pilot study was used to test the possibility of using an online platform for the questionnaire. The use of an online platform is attractive as it provides the advantage of reaching a larger number of potential subjects. The objective of the study was set to determine the effect of different translation procedures on the perception of CVNCs in excerpts from a TV anime on DVD. Perception was measured in terms of PCEs and attention distribution, rather than PCEs and processing effort. Measuring attention distribution entails different methods than measuring processing effort and it can be relatively easily measured by counting the number of fixations in different areas of the image. A relatively simple measurement was used in order to ensure that the main focus remained on the handling of the technology.

The pilot study considered three translation procedures on the basis of the classification used by Chaume Varela (1997) to deal with nonverbal information in dubbing: *equivalence*, *explicitness* and *total substitution*. *Equivalence* occurs when the subtitlers leave the CVNC as it is, *explicitness* when they explain it in some way, and *total substitution* is when the original meaning is altered to fit the target text. Chaume Varela's (ibid) scale was used as it was found to be applicable to nonverbal items, and also because it provided translation procedures at either extreme of the spectrum of interference from the translator (*equivalence* and *total substitution*), as well as a more mid-spectrum procedure (*explicitness*). The experiment was set up to show subjects excerpts from a subtitled TV anime, after which they were asked to rate their level of understanding of a CVNC before answering an open question on it. It was expected that the more explicative the procedure used in translation, the higher the number of subjects who correctly answered questions about the CVNC would be, and that the use of total substitution would lead to a lower number of subjects answering correctly.

7.1.2 Subjects

The subjects for the pilot study were obtained from a selection of students from DCU. A total of 58 subjects volunteered for the study, made up of graduate and undergraduate students. Four undergraduate students from the group took part in the eye-tracking portion of the experiment.

7.1.3 Material

The source material used in the pilot study was the anime *Doki Doki School Hours* (2005). It was first broadcast in Japan in 2004 on TV Tokyo, a channel that broadcasts over half of the weekly animation shows aired on terrestrial television networks in Japan, as well as some of the world's most successful TV anime series such as ポケットモンスター (*poketto monsutā*, Pokémon, 遊戯王 (*yūgiō*, Yu-Gi-Oh!) and ナルト (*naruto*, Naruto) (TV Tokyo Corporation 2006). Volume one of the North American DVD release from 2005, distributed by Geneon was used for the study²⁶. *Doki Doki School Hours* (ibid) revolves around a high school class whose teacher, Suzuki Mika, is in her twenties but

²⁶ It contains the first four episodes of the series, which are based on the first collected volume of the manga of the same name published by Takeshobo in 1999 in Japan.

looks like a child because of her small stature. It consists of short sketches surrounding the teacher and her students, who portray a variety of Japanese high school stereotypes.

The excerpts selected for inclusion in the study were chosen according to their content, following the procedure as detailed. Firstly, 15 excerpts were extracted from the series and shown to a panel of five postgraduate students to seek preliminary feedback. This resulted in reducing the number of excerpts to 10. These excerpts were then digitised for incorporation into the online questionnaire, and sample questions were elaborated before uploading the content to the online platform. However, upon testing the questionnaire, low download speeds because of the large file size of the excerpts indicated the difficulties subjects would have in downloading the files. This eventually resulted in the decision to further decrease the number of excerpts to six in an effort to reduce the amount of time required by the subjects to download the files.

7.1.4 Questionnaire design

The questionnaire, which can be found in Appendix B, was divided into three sections. The first section asked for subjects' biographical information, and some short questions about anime, while the second section consisted of content questions about the excerpts. Each excerpt had two questions; the first of which asked subjects to rate their understanding of the CVNC in question on a Likert scale from zero to six²⁷, and the second of which was an open question to test their actual understanding of the CVNC.

7.1.5 Procedure

The questionnaire was distributed in two forms, online and in print form. In order to maximise the number of responses, several group sessions were organised where the print form questionnaire was distributed to a group of subjects, and the excerpts were projected on a white board using a data projector. The dissemination of the questionnaire on the virtual learning environment (VLE) platform *Moodle*, which is adopted campus-wide in DCU, meant that the subjects could, theoretically, fill it out on any computer with

²⁷Zero represented a complete lack of understanding and six represented total understanding.

internet access. Subjects were given a password to gain access to the secure site and the questionnaire was designed so that each subject could fill it out only once, but could leave and return to it in several sessions. In the end only six subjects took the online form of the questionnaire, and several emails were received where prospective subjects cited reasons for their discontinuation of the online questionnaire. These included slow download speeds for the video files, difficulties viewing the video files on their web browser, and the browser crashing while they completed the questionnaire.

7.1.6 Issues arising from pilot study

Several issues arising from the pilot study highlighted difficulties with using both the eye tracker and VLE platform, as well as potential problems with the design of the experiment. These will now be discussed under technological and experiment design issues.

7.1.6.1 Technological issues

As noted above, despite the easier accessibility, the number of pilot study subjects who answered the internet-based form of the questionnaire was quite low, equalling roughly 10% of the total 58 subjects of the pilot study. The problems encountered by subjects who could not complete the internet-based questionnaire despite their attempts drove the decision to use only paper-based questionnaires in the final experiment. The issue of video file download speeds would have been even more noticeable in the subsequent main experiment than the pilot study, as the number of excerpts had to be increased from six to eleven.

Another technological issue concerned the preparation of the excerpts for use with *ClearView*. As mentioned in Section 5.3, because *ClearView* can only be used with video stimuli in the .avi format, the video had to be converted from the .vob file storage format found on DVD. While it was relatively easy to extract the video and audio from the DVD and convert it to .avi, doing so with the original subtitles intact proved quite difficult. A detailed explanation of the process of excerpt extraction and conversion is provided in Appendix A.

7.1.6.2 Experiment design issues

In addition to the technological reasons, the need for the experiment data to be collected with the questionnaire and eye tracker for all subjects was another factor leading to the dismissal of internet-based questionnaires. The use of the eye tracker made it necessary in any case for subjects to participate in the experiment individually in person. Data analysis also required more time, as several fixation-based measurements not used in the pilot study needed to be used in the main experiment, as well as pupillometric data, which required a comparatively long time to analyse.

ClearView requires that scenes and AOs be defined before many of the functions can be used. In the pilot study, scenes were defined for the duration of time that the CVNC was on screen. While this allowed for a focus to be placed mainly on the CVNC, it meant that comparisons could not easily be made between excerpts, as the CVNCs were on screen for different amounts of time. The definition of scenes was therefore altered in the experiment (see Section 7.6.3.3).

Some of the questions in the questionnaire were also changed in the experiment. As explained previously, in the pilot study subjects were asked to rate their understanding of CVNCs using a Likert scale between zero and six. This resulted in the scale having a mid-point value (three), which would not allow for straightforward categorisation of those subjects who thought they understood and those who did not. For that reason, the scale was altered in the experiment to avoid having a mid-point value.

More data was required for the analysis in the main experiment than the pilot study, so eleven excerpts were used rather than six, and the focus was narrowed to only one translation procedure. As well as the aforementioned reference to the number of subjects in previous studies where eye tracking was used with AV content (see Section 6.5), the anticipated increase in the data to be analysed was another factor behind the decision to reduce the number of subjects in the study, from 58 to 20.

7.2 Hypothesis, independent and dependent variables

The present study aims to examine the correlation between the presence of pop-up gloss and viewer perception of subtitled excerpts from a TV anime DVD in terms of PCEs and processing effort experienced by the subjects.

The analysis of different forms of data (questionnaire, fixation-based and pupillometric) in the present study required that separate hypotheses be used with each data set. The hypotheses tested with questionnaire and eye tracker data are explained in detail in Sections 8.1.1 and 8.2.1 respectively.

The independent and dependent variables used with consecutively and simultaneously collected data also differed, and are discussed in the following.

7.2.1 Questionnaire data: Independent variables

Independent variables can also be referred to as change variables, as they are responsible for bringing about a change in a phenomenon or situation (Kumar 2005). There was only one independent variable used with the questionnaire data. This was whether the pop-up gloss procedure was used in an excerpt or not, resulting in two experiment conditions, *treatment* and *normal*. The treatment and normal conditions contain “abusive” and “corrupt” subtitling strategies respectively, under Nornes’ (1999) terminology (see Section 2.1). The treatment condition uses an experimental procedure, i.e. “abusive” subtitling, while the normal condition adheres to subtitling norms, i.e. “corrupt” subtitling. Examples of screen captures from both conditions are shown in Figure 7.1.



Figure 7.1. Screen capture from the same scene in *treatment condition* (left) and *normal condition* (right) (*Paniponi Dash!* 2007)

7.2.2 Questionnaire data: Dependent variables

Dependent variables can be described as “the outcome of the change(s) brought about by the introduction of an independent variable” (Kumar 2005 p60). The questionnaire data was used to compare several dependent variables. The comparison of these variables was made between subjects who saw the excerpts in the treatment condition and normal condition. The variables that were used in this study were different for each of the aspects of perception that were measured: PCEs and processing effort.

i.) PCEs:

PCEs were measured in terms of *actual* PCEs, calculated using the number of correct answers by the subjects to open questions on the subtitle and CVNC; and *declared* PCEs for the CVNC, measured using the Likert scale provided in the question on the CVNC. Likert scores between 4 and 6 were categorised as “declared understood”, and those between 1 and 3 as “not understood”.

ii.) Processing effort:

The processing effort required for each excerpt was measured in the questionnaire in terms of the ease with which subjects could read the subtitles. To determine the ease with which subjects could read the subtitles, they were asked to rate the perceived speed of the subtitles. Instances where the subtitle speed was rated on the Likert scale between 4 and 5 were classed as “declared fast”.

7.2.3 Fixation-based and pupillometric data: Independent variables

The fixation-based and pupillometric data required the use of two independent variables, or factors, i.e. *display type* and *subtitle length*, each with two levels. The two levels of the factor *display type* were “gloss” and “solo” (“gloss” when a pop-up gloss appeared on screen with a subtitle, and “solo” when a subtitle appeared on screen alone), and the two levels of the factor *subtitle length* were “one line” and “two line”. The factor of subtitle length was introduced as significant differences between the viewing of one line and two-line subtitles have already been observed and accepted (d’Ydewalle and Gielen 1992, d’Ydewalle and de Bruycker 2007). Both factors were considered as within-subject variables, and were crossed resulting in altogether four experiment conditions to which each subject was exposed, as illustrated in Table 7.1. Screen captures showing an example from each of the conditions can be seen in Figure 7.2.

<i>Experiment condition</i>	<i>Display type</i>	<i>Subtitle length</i>
1	Solo	One-line
2	Solo	Two-line
3	Gloss	One-line
4	Gloss	Two-line

Table 7.1. Experiment conditions used with fixation-based and pupillometric data



Figure 7.2. Screen captures depicting an example from each experiment condition described in Table 7.1 (*Paniponi Dash!* 2007)

7.2.4 Fixation-based and pupillometric data: Dependent variables

Four dependent variables were used with the fixation-based data in the present study, all of which have been used in perception studies of subtitled content in the field of psychology (d'Ydewalle & de Bruycker 2007):

1. The *percentage of skipped subtitles* was calculated by counting the number of subtitles that were not fixated upon at least once and dividing it by the total number of subtitles. For example, if there were 50 subtitles in one condition and 25 of them were not fixated upon by one subject, the percentage of skipped subtitles for that subject in that condition would be 50%.
2. To obtain the *percentage of gaze time in subtitles*, the percentage of time spent in the subtitle area while a subtitle was displayed on screen was calculated. For example, if a subtitle was displayed on screen for two seconds, and the subject fixated on the subtitle area for a total 0.5

seconds of that time, the percentage gaze time would be $0.5/2 = 0.25$ (x 100%) = 25% gaze time.

3. The *mean fixation duration* was obtained by adding the duration of all fixations in the subtitle area and dividing this figure by the number of fixations in the area. So if there were five fixations on a subtitle, and the total gaze time on the same subtitle was 500 ms, the mean fixation duration would be 100ms.
4. The *word fixation probability* was calculated by dividing the total number of fixations in the subtitle area by the number of words in the subtitle. So a 12-word subtitle that was fixated upon six times would have had a word fixation probability of 0.5. This weighting was applied because the average length of subtitles in the analysis was not the same for each subject. This is only an approximate measure as the number of words that are skipped or refixated influence the value of the variable (d'Ydewalle & de Bruycker 2007 p199).

The mean fixation duration and word fixation probability were only calculated for subtitles that were fixated upon at least once, while the percentage of skipped subtitles and the percentage of gaze time in subtitles were calculated for all of the subtitles.

The dependent variable used with pupillometric data was *median pupil size*, which was calculated as follows. First of all, the pupil size data for the left and right pupils was exported to a spreadsheet along with the time data for the pupil recording. The original video files were then used to measure the in and out times of the subtitles, which were used in conjunction with the time data for the pupil recording to note the times when subtitles were on screen. A mean value for both pupils was then calculated in an effort to counter any missing or corrupt data from one pupil. In a further effort to eliminate outliers, such as data variation caused by blinking²⁸, a median value was taken for each subject and subtitle rather than the mean. The mean data for each subject and each of the

²⁸ When the eyelid is closed, pupil data cannot be gathered, and when the eyelid opens after blinking, there is a brief pupillary reaction to light, not caused by cognitive effort (Schultheis 2004).

fixation-based and pupillometric dependent variables can be found in Appendix C.

7.3 Units of analysis

The unit of analysis used in the experiment in the present study was dependent on the type of data collection used. Howell's (2005) distinction between translation procedures, or techniques used to deal with individual textual items, and translation strategies, as the application of these procedures throughout a text, will be applied here.

Data gathered consecutively using the questionnaire provided information on the relationship between the presence of pop-up gloss and the PCEs and processing effort as experienced by viewers in the excerpt as a whole, and thus focused on the translation strategy used in the excerpts. This data was used to compare subjects who saw the same excerpts in different conditions.

For simultaneously gathered data, a different unit was required to analyse the results. While the questionnaire data was compared depending on the presence of pop-up gloss in the excerpt, simultaneously gathered data was considered in terms of individual subtitles, focussing on the translation procedure being used within the excerpts. This resulted in a within-subject comparison being carried out, where the values for each subject were compared individually, for each experiment condition.

7.4 Statistical testing

The data gathered from both the questionnaire and the eye tracker was analysed using statistical tests. Statistical tests allow the researcher to estimate whether the results from a particular data sample provide enough evidence to prove a hypothesis on a larger scale. They are usually used to generalise a result for an entire population. However, as the sample size in the present study was quite small, the results of the statistical tests should be regarded conservatively. Nonetheless, they are still a valuable method of analysing the trends in the results and provide a relatively objective benchmark for determining whether the difference in data values is significant or purely

because of chance. Conyngham (2007) suggests four steps that should be taken in the statistical testing process:

1. Stating the null and alternative hypothesis:

The null and alternative hypothesis are opposing statements about the population parameter being tested. In general, the researcher tries to prove the alternative hypothesis, by showing that the null hypothesis is false. The null hypothesis is represented with H_0 and the alternative hypothesis with H_1 . So, for example, if the researcher wanted to suggest that the average pupil size is larger when there is a pop-up gloss present, it could be represented in the following way.

First of all, the null hypothesis is stated, that is, the hypothesis the researcher wants to disprove. In this case the researcher wants to show that average pupil size in the gloss condition is larger than in the solo condition, and so will attempt to rule out the other possibilities (i.e. that it is the same or smaller in the gloss condition) found in the null hypothesis. For the sake of brevity, the null hypothesis can be represented using this formula:

$$H_0 : \text{Average pupil size}_{\text{gloss}} \leq \text{Average pupil size}_{\text{solo}}$$

Next, the hypothesis the researcher is trying to prove, the alternative hypothesis, is stated. In the example here, it could be represented by the formula:

$$H_1 : \text{Average pupil size}_{\text{gloss}} > \text{Average pupil size}_{\text{solo}}$$

2. Choose the appropriate statistical test:

To do this, the researcher must decide what the test is supposed to prove and the type of variable being used. This is discussed in further detail in the following section.

3. Calculate the test statistic

In the present study, this was done using the freely available software packages *OpenStat*²⁹ and *ezANOVA*³⁰. T-tests and the Shapiro Wilk *W* test

²⁹ Available for download at <http://www.statpages.org/miller/openstat/OpenStat.zip>, developed by William G. Miller at Iowa State University

were carried out using the statistical analysis software, *OpenStat*. Analysis of variance (ANOVA) tests were calculated using another piece of statistical analysis software, *ezANOVA*.

4. Interpret the test statistic/p-value:

While several test statistics may be used depending on the statistical test being applied, there is a measure called the p-value, or the probability value, which is common to all statistical tests. The p-value represents the probability of an outcome occurring by chance. The decision to accept or reject the null hypothesis is based on the p-value. If the p-value is less than 0.01 (e.g. $p=0.009$), for example, there is a less than 1% chance that the null hypothesis occurred by chance, and thus a more than 99% probability that it is false.

In order to determine the value at which a p-value is going to be considered statistically significant, the α -level must be specified. The α -level is an arbitrary value but it is usually set at 0.05, 0.01 or, less frequently, 0.001 (Lang 2007). The significance level used in the present study for all statistical tests is $\alpha=0.05$, which means that a p-value had to be less than 0.05 for it to be considered significant. This α -level was chosen because it is the level of significance most frequently used (Lang *ibid*). P-values are reported here in one of four categories, $p<0.05$, $p<0.01$, $p<0.001$, each representing increasingly statistically significant results, and $p>0.05$, representing non-statistically significant results.

The interpretation of the test statistics is reported in the data analysis in Chapter 8. In the following, the statistical tests used and the motivation behind their selection is discussed.

7.4.1 Statistical test used with questionnaire data: The paired sample t-test

The statistical test selected for use with the consecutively collected data in this experiment was the paired sample two-tailed t-test. The t-test is one of the

³⁰ Available for download at <http://www.sph.sc.edu/comd/rorden/ezanova/ezanova.zip>, developed by Professor Chris Rorden at the University of Southern California

most widely used tests to compare the differences between the means of two samples with only one independent variable (Miller 2008). The paired sample t-test is used when each member of one sample has a unique relationship with a particular member of the other sample (Fadem 2008), as is the case in the current experiment where a comparison is made between the means of a variety of dependent variables (see Section 7.2.2) for subjects who saw the same excerpt in one of two conditions.

The test statistic provided by the t-test is a t-value, which can in turn be used to calculate a p-value. As paired t-tests belong to a group of tests termed parametric tests, they are based on the assumption that the distribution of the variable being tested is normal (Miller *ibid*). However, it can be difficult to determine when the data sample is small (<50), so in the present study, a normality test was used to determine the likelihood that the questionnaire data came from a normally distributed population of values.

The Shapiro Wilk W test (Shapiro and Wilk 1965) is a normality test used to check whether a data sample came from a normally distributed population, and the test statistic it provides is a W value. If the W value is significant ($p < 0.05$), there is a higher likelihood that the data does not come from a normal distribution. As the results for each data set in Table 7.2 show, there was no statistical evidence against the normal distribution of all the data used in the analyses above, so it was considered appropriate to use parametric testing.

<i>Data</i>	<i>Shapiro Wilk W test statistic/p value</i>
CNVC actual understanding – treatment condition	W=0.9592, p>0.05
CNVC actual understanding – normal condition	W=0.9496, p>0.05
Subtitle actual understanding – treatment condition	W=0.9603, p>0.05
Subtitle actual understanding – normal condition	W=0.9160, p>0.05
CVNC perceived understanding – treatment condition	W=0.9143, p>0.05
CVNC perceived understanding – normal condition	W=0.8917, p>0.05
Subtitle perceived fast – treatment condition	W=0.8637, p>0.05
Subtitle perceived fast – normal condition	W=0.9498, p>0.05

Table 7.2. Results of Shapiro Wilk W tests for normal distribution of questionnaire data

As a further means of controlling for the relatively small sample size, a two-tailed t-test was chosen over a one-tail t-test. This is because “two-tailed t-tests require a greater difference to produce the same level of statistical significance (i.e. p-value) as one-tailed tests. They are more conservative and often preferred for this reason” (Lang 2007 p318).

7.4.2 Statistical test used with fixation-based and pupillometric data: Two factor within-subject analysis of variance (ANOVA)

As there were two independent variables used with the fixation-based and pupillometric data, the t-test was no longer appropriate. Instead the data gathered with the eye tracker was analysed using a different parametric statistical test, the analysis of variance, or ANOVA. ANOVA tests are used to analyse the variability of test objects and the test statistic they provide is an *F* value (Salkind 2008). There are several types of ANOVA, and the appropriate test is chosen depending on a number of factors such as how subjects are grouped and if the same variable is measured multiple times for each subject. The ANOVA test is a popular statistical test in studies of subtitled media using fixation-based data (e.g. de Linde and Kay 1999, d’Ydewalle and de Bruycker 2007), and studies where pupillometric measurements were used (Just and Carpenter 1993, Schultheis 2004, Schultheis and Jameson 2004). It allows the researcher to measure the interaction of multiple independent and dependent variables. Separate ANOVA tests were carried out for each of the dependent variables measured (see Section 7.2.4).

The type of ANOVA that was used in the present study is a within-subject ANOVA, also known as a repeated measures design, which compares the means of dependent variables for the same subjects in all experiment conditions. As well as deciding on either between-subject or within-subject ANOVA, the type of ANOVA calculation depends on the number of factors and levels of each factor in an experiment. A factor is a category of independent variable, so in the present study there were two factors, i.e. subtitle length and display type. Each factor must have at least two levels, which is a setting of the factor. For example, in the current experiment, both factors have two levels: the factor subtitle length has the levels one-line subtitle and two-line subtitle; and the factor display type has the levels gloss and solo. This results in what is called a 2 x 2 factorial design, as there are two factors with two levels each. ANOVA tests where there is more than one factor not only allow us to look at the individual effects of each factor, but also to look at the simultaneous effects of both, called an interaction (Salkind 2008).

7.5 Experiment validity

On the basis of the lessons learned from the pilot study, the hypothesis, variables and statistical tests that would be used in the experiment were finalised, before devising a revised experiment design. The revised design was then considered under Frey *et al.*'s (1991) three categories of experiment validity: measurement, internal and external validity. They are discussed in the following.

7.5.1 Measurement validity

Measurement validity is differentiated from measurement reliability, with validity determining the extent to which a measurement actually relates to the concept it is supposed to measure, while reliability refers to the extent that the measure is consistent and repeatable (Lee 2006a).

One way of ensuring measurement reliability is to use triangulation. Triangulation involves the application of several instruments of data gathering and analysis to investigate the nature of a phenomenon (Alves 2003). To improve the intersubjectivity of the data, triangulation was used in the present

study, with a combination of eye fixation data, pupillometric data and questionnaire data. The reliability of these forms of data as measures of PCEs or processing effort is supported by their being drawn from prior research discussed. The notable exception is pupillometric data, as its validity as a measure of processing effort experienced when watching AV content has not previously been established. However, the use of triangulation offered a means of comparing and validating the pupillometric data using fixation-based and, to a lesser extent, questionnaire data.

7.5.2 Internal validity

The internal validity of a study refers to the accuracy of the results based on its design and execution, and it can be influenced by several factors, including measurement validity and reliability, and data collection procedures. Measurement validity and reliability were dealt with above, so the following will begin with a description of how data collection procedures can influence internal validity, before describing the effects of some other threats to internal validity.

7.5.2.1 Data collection procedures

According to Lee (2006b) two prime aspects must be attended to when collecting data, i.) the study setting and procedure should be standardised, and ii.) valid manipulations and treatments should be used. In order to ensure the first of these, an experiment protocol was developed, in which the majority of information was presented to subjects in written form. All of the subjects took part in the study in the same room, sitting at the same workstation, with the same volume level on the speakers. Lighting levels were also kept relatively constant, by closing the blinds in the lab and turning on the same one light for all subjects. The relative stability of the lighting levels was a particularly important factor in this experiment, because of the sensitivity of the pupillary response to light.

The validity of working manipulations and treatments used was dealt with in two categories, depending on the data being analysed. In the analysis of data gathered using the questionnaire, comparisons were made between abusive and corrupt strategies (Nornes 1999), so the excerpts that subjects watched

were identical except for the subtitle strategy that was employed, resulting in a between-subject comparison.

As the data gathered using the eye tracker was measured simultaneously, it was therefore possible to link it to individual subtitles. As each subject watched a combination of excerpts with and without pop-up gloss, a within-subject comparison was considered to be more appropriate here than a between-subjects comparison. However, the presence of uncontrolled content in the other visual nonverbal, audio verbal and audio nonverbal semiotic channels raised concerns that the pupillometric measurements may have been affected. It was hoped that the volume of subtitles analysed (at 202 in total, see Table 8.4) with content varying in these channels would counter this. Also, in other eye tracking studies where subtitled AV content was used (de Linde and Kay 1999, d'Ydewalle and de Bruycker 2007), a similar situation is present, although these admittedly do not use pupillometric measurements.

The particular menu system used on the experiment material, the *Paniponi Dash!* (2007) DVD, was a further complicating factor in the study. The way the DVD extras are organised means that the pop-up notes can only be viewed along with the “full caption subtitle track”. This track included extra text on screen while some subtitles were on screen, and may have affected eye behaviour in subjects who viewed the excerpt in the treatment condition. Figure 7.3 shows an example of a screen capture from the same point in an excerpt with the full caption subtitle track and normal subtitle track, with the extra text highlighted with red boxes.

The low percentage of subtitles for which this extra text was present (6.93%³¹) was hoped to limit the influence this had on the results. This lack of control over all variables, although undesirable, is often an unfortunate circumstance when the product to be analysed is a commercial release, such as is the case in the current experiment.

³¹ This figure was calculated by dividing the number of subtitles on screen at the same time as extra text over the total number of subtitles and multiplying the resulting figure by 100.

7.5.2.2 Maturation

This measure refers to the physical and psychological changes that occur in subjects naturally over time, and that are not a result of variables in the experiment. This tends to be more problematic for longitudinal studies, i.e. studies that span over a long period of time. As all of the data collected for this study was gathered over a period of two weeks, and over an hour for individual subjects, this factor was not expected to pose a threat to the data in the current experiment.



Figure 7.3. Screen capture from the same point of an excerpt with full captions (top) and normal subtitles (bottom) (*Paniponi Dash!* 2007)

7.5.2.3 Attrition

Attrition refers to the loss of aspects of the study that were present at the beginning as the study progresses. Subjects were compensated for taking part

in the study³², perhaps lending to the lack of anyone dropping out of the study. In addition, the moderate length of the study (45 minutes approx.) was set in order to avoid fatigue-related drop out. While all of the questionnaire answers were valid, a lower limit of eye track data which had to be present for the subject to be included in the fixation based and pupillometric analyses was not met by six of the subjects. This resulted in their data being subsequently excluded from the eye-tracking portion of the data analysis. One further subject was excluded from the pupillometric analysis as she reported a medical condition where her pupils do not react normally to light.

Several factors may have hampered the collection of data with the eye tracker, such as blinks and the head being too far from the camera. Table 7.3 shows the percentage of data present for each subject and each excerpt. This was calculated by copying the raw pupil dilation measurements into an *Excel* spreadsheet and deleting all instances where a -1.00 measurement was present. The -1.00 measurement indicates that there was no valid data recorded for the eye. After all these instances were deleted, the number of cells that contained data were counted and divided by the total number of cells. This figure was calculated for the left and right eye, and the percentage of both was averaged to give the overall percentage of missing data. The numbers in bold italics indicate all instances where less than 85% of the data was present. In order to keep the comparisons more even, particularly concerning the pupil data that showed a large degree of variance in both the size and degree of the change between subjects, it was decided to exclude the subjects who did not have at least 85% of pupil data present in all excerpts. This resulted in the exclusion of the six subjects highlighted in grey in Table 7.3.

³² The level of compensation was not unusually high, and was included in the proposal to the DCU Ethics Committee, which was subsequently approved.

Excerpt no.	1	2	3	4	5	6	7	8	9	10	11
Subject no.											
1	90.13	88.97	95.27	95.37	90.83	94.78	94.74	91.19	90.09	95.04	94.39
2	24.20	65.13	55.15	49.37	52.00	57.61	71.87	73.30	55.41	74.67	64.83
3	98.89	98.89	98.01	98.17	99.31	98.43	95.57	98.74	95.31	98.16	97.80
4	99.46	99.33	99.00	98.70	99.61	99.05	99.27	99.49	99.23	99.30	98.86
5	71.68	74.55	58.83	47.84	69.08	57.77	78.63	73.77	60.25	72.32	60.35
6	99.46	99.65	99.46	99.26	99.28	99.51	98.51	99.55	94.90	97.99	98.10
7	99.30	98.68	98.00	97.60	97.71	97.12	96.86	95.31	97.19	98.64	97.99
8	94.36	96.01	95.80	91.40	94.61	91.35	94.18	93.55	90.59	94.10	89.60
9	81.78	98.49	96.18	45.57	3.01	92.52	73.92	44.85	81.04	58.26	80.42
10	89.58	85.37	86.80	88.19	91.07	90.71	86.70	87.54	95.47	87.77	93.83
11	97.33	94.15	94.48	91.55	90.78	94.89	96.45	96.66	97.05	95.46	97.86
12	96.03	97.05	96.18	96.86	97.16	93.87	95.93	98.75	91.78	97.44	95.94
13	98.86	98.55	98.98	96.30	97.03	99.09	99.70	98.88	97.48	99.28	98.57
14	96.99	95.50	97.55	97.53	97.58	96.70	95.97	97.52	97.23	95.24	95.47
15	53.08	13.81	22.10	23.06	24.66	13.25	25.85	15.53	18.85	22.52	13.31
16	98.58	99.51	97.94	99.33	99.26	99.25	99.52	98.69	97.76	99.19	97.93
17	99.58	98.78	98.17	97.35	92.14	97.11	99.09	96.62	98.44	99.64	95.56
18	50.04	51.06	42.16	46.95	41.69	47.87	49.80	39.44	32.80	42.19	71.37
19	98.31	99.39	99.95	97.50	98.73	99.42	97.89	99.21	97.09	99.18	98.38
20	80.36	88.64	96.43	96.17	95.89	95.50	65.01	96.60	95.26	97.76	94.79

Table 7.3. Percentage eye data present for each subject and excerpt

7.5.2.4 Sensitisation

Sensitisation occurs when subjects in the study are familiar, or become familiar, with the research procedures. The design of the experiment meant that subjects were asked questions relating to the same four aspects for each excerpt (PCEs relating to the CVNC, declared PCEs relating to the CVNC, processing effort and PCEs relating to subtitles). While the subjects were deliberately kept naïve to the true purpose of the experiment and only told the broad purpose of the study, there was some concern that they would realise that the answer to the CVNC question could be found in the pop-up gloss and begin to pay a disproportionate amount of time viewing the gloss. Consequently, the subjects were shown the excerpts alternating between conditions where pop-up gloss was used and where it was not used. It was hoped that this would keep subjects naïve to the fact that the CVNC answer was in the pop-up gloss. This is because the lack of pop-up gloss in some excerpts would mean that subjects would remain unaware of the presence of the answer to the CVNC question in the pop-up gloss. After filling out the

questionnaire, one subject, whose data was disregarded anyway due to his not meeting the 85% threshold of pupil data being present, did note that he spent more time looking at the pop-up area as the experiment progressed because he had noticed that there was always an answer to one of the questions in the pop-up gloss.

In order to check whether this was the case for other subjects, the percentage gaze time spent in the pop-up gloss area was calculated for subjects 1-10 whose eye data was used in subsequent analysis, for the first 30 instances when a pop-up gloss was on screen. As we can see from the shape of the curve in Figure 7.4, which displays the percentage time spent in the pop-up gloss area in the order in which they occurred, there is only a very weak positive correlation between the percentage gaze time spent in the pop-up gloss and the order in which the pop-up gloss appeared ($r(28)=0.15$). This weak correlation suggests that the subjects did not spend more time in the pop-up gloss as the study progressed, and that the percentage time spent in the area varied throughout the study.

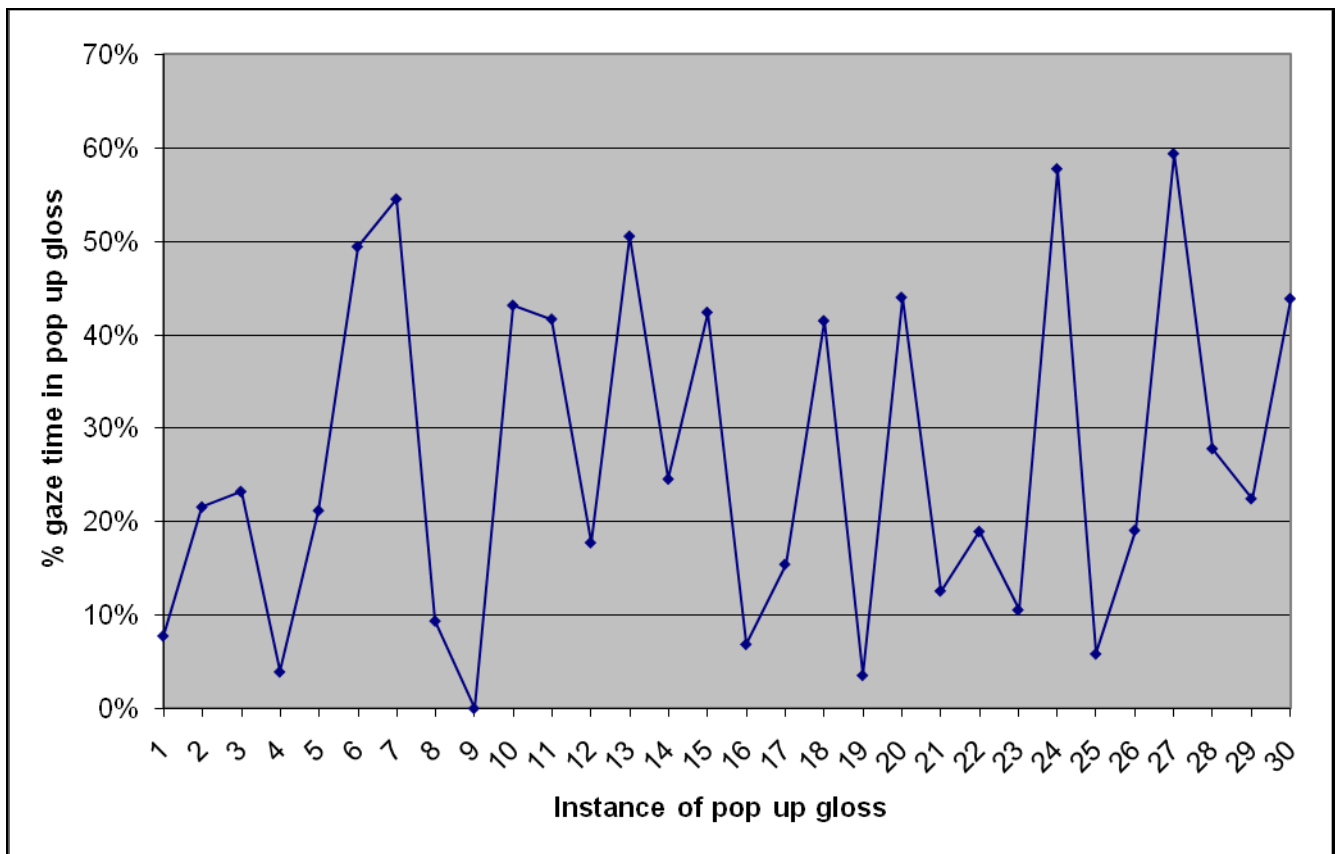


Figure 7.4. Percentage gaze time in pop-up gloss area in order of appearance of pop-up gloss for selected subjects

7.5.2.5 Inter-subject interaction

This threat to validity is caused by the interaction of subjects who may affect each other's answers or results by doing things such as discussing the experiment. The fact that eye tracking may only be carried out on one subject at a time meant that the threats because of inter-subject interaction were minimal. However, the design of the study necessitated that the subjects remained naïve to the real purpose of the experiment in order to emulate as closely as possible "genuine" eye movements of viewing anime. For this reason, subjects were given individual time slots for the experiment, so that they would not congregate outside the lab and have the opportunity to discuss the experiment, and were asked not to discuss the real purpose of the experiment with others who were going to participate in the study.

7.5.2.6 Hawthorne effect

This is caused when subjects behave differently in an experiment because they know they are being observed (Lee 2006b). Of course, the fact that subjects were aware their eye movements were being recorded meant that they may have been aware of their eye movements at first. However, the fast moving pace of the excerpts, the task of answering questions in the questionnaire, and the unobtrusive nature of the eye tracking monitor were all hoped to have made the subjects adjust to the situation more quickly and forget that their eye movements were being recorded by the eye tracker.

7.5.2.7 Researcher presence

This threat can occur because of the researcher's presence affecting the data collected, or by the researcher inadvertently revealing the type of data they desire to the subjects and causing them to change their behaviour. This could theoretically be avoided by hiring research assistants to carry out the experiments. However, the time needed to train them in using the eye tracker, and incurred expenses for compensating them meant that this was not feasible, so instead a protocol was used, which is described later (see Section 7.6.4).

7.5.3 External validity

The external validity of a study is concerned with how accurately its findings can be generalised. This may be influenced by factors such as sampling, ecological isomorphism and replication. The issue of sampling is dealt with in the discussion about subjects in Section 7.6.3.4.

7.5.3.1 Ecological isomorphism

This refers to the proximity of the setting of the experiment and the real world setting. In the present study, while it would have been desirable to allow the subjects to view the excerpts on their own televisions at home, this was not feasible. However, the general trend in the increased use of both desktop and laptop sized monitors for watching AV content was evident in the subject group, with all of them having watched AV content on a PC monitor and all bar two having watched subtitled content on a PC monitor. The workstation that was used in the laboratory had a comfortable chair for subjects and the lighting of

the room and volume of the speakers were both set at a comfortable level. The monitor-integrated eye tracker has an advantage over head-mounted eye trackers or those where the chin must be placed on a rest in regards to ecological isomorphism, as it is less obtrusive and allows for head movement and a more realistic viewing environment.

Subjects' inability to pause the excerpts was another drawback resulting from the use of the eye tracker. Several subjects raised this issue after they viewed the excerpt in the treatment condition, with pop-up gloss. One subject, when asked whether he would watch a whole DVD with pop-up gloss answered he would, but with "a lot of pausing to read the notes" (Subject 3), while another answered "I would have to pause the DVD to read them" (Subject 15). This is one of the main weaknesses of the experiment design, but was unavoidable if the eye tracker was to be used. The issue is discussed further in the conclusion under the strengths and limitations of the thesis (see Section 9.2).

7.5.3.2 Replication

Replication refers to the verification of the results of the present study by comparing them to similar studies. The fixation-based measurements used in the current experiment were used in previous studies (e.g. d'Ydewalle and de Bruycker 2007, d'Ydewalle and Gielen 1992) to compare one and two-line subtitle conditions. The inclusion of fixation-based measurements that have been used before provides a standard to which the results obtained here may be compared. The same is not true for the pupillometry results, where no prior study has used the pupil size as a measure of cognitive load with subtitled AV content. For this reason, the design of the experiment is outlined in detail to help ensure its replicability.

7.6 Experiment design

The experiment was designed to measure the effect that the pop-up gloss has on viewer perception of subtitled TV anime on DVD. The two aspects of the TV anime text on which the current experiment focused were subtitles and CVNCs. Perception was considered from a relevance theoretical perspective, and it was analysed in terms of i.) PCEs and ii.) processing effort. The design of the experiment and procedure carried out are described in the following.

7.6.1 Source text selection

The choice between using a commercially produced DVD and using fan-subtitled content from the internet was a factor that had to be considered before selecting the TV anime to use in the experiment. In the end a commercially produced DVD was selected in order to enable a fairer analysis to be made between the excerpts. All excerpts came from one DVD and were processed in the same way, while if online content were to be used it would be more difficult to verify whether the processing of various online episodes would have been uniform. Factors such as sound and graphic compression rates were expected to vary less if the excerpts were all taken from one DVD, rather than if two versions of the same excerpt were downloaded from fan sites on the internet, where various compression formats used could affect the sound or image adversely.

The series used was the first DVD volume of the TV anime series *Paniponi dash!* (2007), distributed in Ireland, the UK and North America by *ADV films*. The series is relatively unknown in Europe but is quite popular in Japan, as is evinced by its ranking in third place (5431 votes) in an online survey on people's favourite hundred anime³³. The primary reason for selecting the series for the experiment was that the DVD offers the option to watch the content with pop-up gloss. As explained in Chapter 2, the pop-up gloss is used in the DVD to explain culturally marked items in each of the semiotic channels, which the examples in Table 7.4 and Figure 7.5 illustrate.

³³ Survey carried out by the TV Asahi program 月バラ (*getsu bara*, Monday Variety) between October 2005 and September 2006 (<http://www.tv-asahi.co.jp/anime100/>).

<i>Semiotic channel</i>	<i>Pop-up gloss</i>
1. Visual nonverbal	People popping out of a giant apple this way is an iconic image from the anime/manga <i>Sazae-san</i> .
2. Visual verbal	“Peach moon” is written on the fan.
3. Audio verbal	She’s saying the kinds of things you say to calm a horse in Japanese
4. Audio nonverbal	Alien Captain’s Japanese voice actor also played Captain Picard in the Japanese dub of a certain famous show.

Table 7.4. Pop-up gloss in each of the semiotic channels

Paniponi Dash! (2007) is a high school comedy in the *seinen* anime genre, revolving around two classes in a Japanese high school. The main character is the teacher, Rebecca Miyamoto, an 11 year-old university graduate, who is accompanied by a maudlin rabbit and spied upon by aliens from time to time. The series contains a lot of parody and a large number of culturally marked items, many referring to Japanese and some to Western popular culture.



Figure 7.5. Examples referred to in Table 7.4 of pop-up gloss explaining culturally marked items in each semiotic channel (*Paniponi Dash!* 2007)

7.6.2 Ethical approval

The awareness to seek ethical approval is increasingly important and becoming an essential part of research protocol in humanities studies. As the experiment involved testing with human subjects, and they would be asked private questions concerning their health in relation to eye tracker requirements (see Section 7.6.3.4), it was necessary to gain ethical approval for the experiment. The ethical approval process involved filling out an ethics review form that outlined the experiment, and sending it to the DCU Ethics Committee. This was sent along with copies of the plain language statement and informed consent forms that would be used in the experiment. The experiment was categorised as a low-risk social experiment and received ethical approval.

7.6.3 Method

In the following, the material and tools used in the experiment are described, and the procedure that was followed is outlined.

7.6.3.1 Material

The *Paniponi Dash!* (2007) DVD contains five episodes, each 25 minutes long. Excerpts were used in the experiment for several reasons. If an entire episode was used, there was a higher chance that fatigue would begin to affect the eye data readings of participants. Furthermore, the content on the DVD was well suited to being split into excerpts, as the storyline is made up of many sub-plots.

To assist in the replicability of the experiment, the initial idea was to select excerpts at eight and 16-minute intervals from each of the episodes on the DVD. However, it was noted that although the high level of culturally marked visual content in this DVD would allow for such a random selection process, this procedure would not be replicable in many other DVDs. Instead, on the basis of Caffrey (2008a), 20 excerpts were shown initially to two volunteer DCU graduates. After viewing the excerpts, the volunteers were asked questions and a discussion was held to pre-test the suitability of the content and number of excerpts. The pre-test revealed that the volunteers had difficulties in understanding some of the questions, and were concerned with the number of excerpts, with both volunteers experiencing significant fatigue.

On the basis of the pre-test feedback, the number of excerpts was reduced to 11 to counter possible fatigue effects, which were a particular concern due to their effects on pupillometric measurements, with the use of too many stimuli in an experiment possibly resulting in a decrease in pupil diameter (Hess 1972).

Although the primary focus of the present study is the perception of Japanese CVNCs, two nonverbal references to Western pop-culture genres were included to allow for comparison with Japanese items in terms of the number of subjects who experienced PCEs. A still image of the CVNC that was the topic of the question in each excerpt is provided in Figure 7.6, and Table 7.5 shows the text from the pop-up gloss(es) describing the CVNC in the treatment condition.

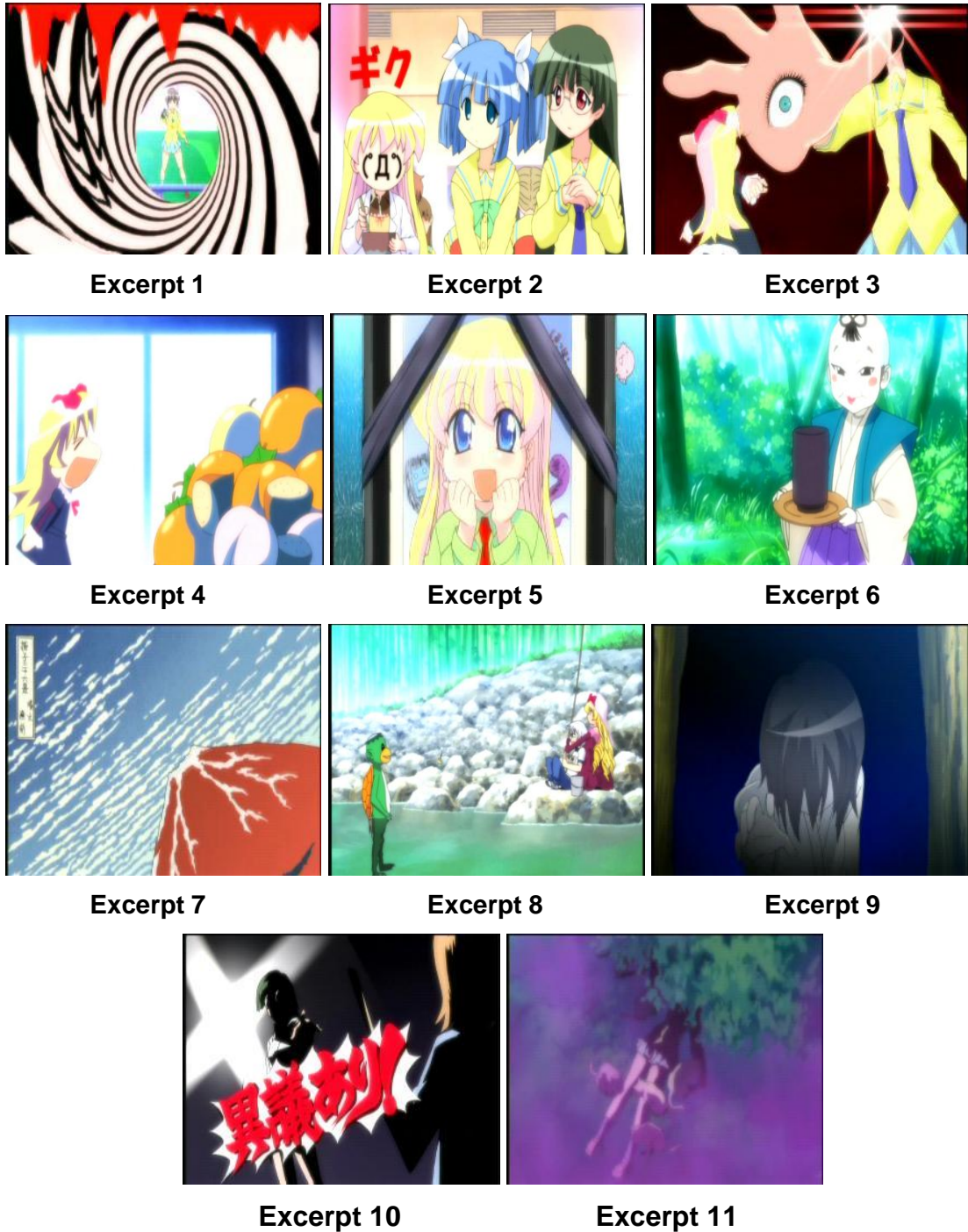


Figure 7.6. Screen captures of CVNCs from each excerpt (*Paniponi Dash!* 2007)

Excerpt no. and CVNC used	Pop-up gloss(es) in treatment condition
<i>Excerpt 1</i> James Bond reference	This is an obvious spoof of the opening sequences to the James Bond movies.
<i>Excerpt 2</i> 2-chan ³⁴ emoticon	2-chan emoticon for shock.
<i>Excerpt 3</i> Eye in palm of hand	<p>There are various Japanese cultural sources for eyes in the palms of someone's hands.</p> <p>In the anime <i>Gegege no Kitaro</i> a scary demon called "Te no Me" ("Hand Eyes") lived up to his name.</p> <p>On the "good guy" side of the fence, Kannon the bodhisattva is depicted with eyes in the palms of each of her 1000 hands.</p>
<i>Excerpt 4</i> Fruit falling	<i>Momo kuri sannen kaki hachinen</i> is a saying: "It takes three years for peach or chestnut trees to bear fruit, and eight for persimmon trees."
<i>Excerpt 5</i> Photo frame with black ribbon	In Japan, the photo frame for a deceased person's picture is decorated with a black ribbon.
<i>Excerpt 6</i> Karakuri doll	<p>This is a karakuri doll.</p> <p>The one here is the classic form designed during the Edo period to serve tea. An early form of robotics!</p>
<i>Excerpt 7</i> Hokusai painting	This picture is a parody of Hokusai Katsushita's famous series of woodcuts, "Thirty-six Views of Mt. Fuji".
<i>Excerpt 8</i> Kappa	<p>Kappa are mythical water imps found in Japanese folklore.</p> <p>Some have said they're actually cryptids, because sightings have been reported.</p> <p>In Shintoism they are considered to be one of the many types of water deities.</p>
<i>Excerpt 9</i> White kimono	Ichijo is sporting the white kimono worn by the dead here.
<i>Excerpt 10</i> Phoenix Wright reference	A reference to the video game known in the US as <i>Phoenix Wright: Ace Attorney</i> .
<i>Excerpt 11</i> Catwoman reference	A spoof of the movie <i>Batman Returns</i> . This is the scene where Catwoman is born.

Table 7.5. Screen capture and pop-up gloss description of CVNC on which CVNC question was based for each excerpt (*Paniponi Dash!* 2007)

7.6.3.2 Preparation of material

The process of extracting the excerpts complete with their original subtitles presented some technical difficulties. A workaround for the conversion of the .vob files on the DVD to .avi files was devised, and is described in Appendix A.

³⁴ 2-chan is a contraction of the name 2-channel, an online chat forum where similar emoticons are used (<http://2ch.net/>).

The file conversion was necessary because *ClearView* only allows video files in the .avi format to be used as video stimuli. This unavoidable process led to some degradation of the video resolution with some dropped frames, which may have had a minor effect on subtitle quality. This, as well as the difference of a few seconds between the excerpt lengths for each group meant that subtitle durations needed to be calculated separately for each group.

The resolution of the extracted excerpts was relatively low (720 x 576 pixels) when compared with the maximum 1280 x 1024 pixel resolution of the *Tobii* monitor. The desktop settings were altered to 800 x 600 pixels in an effort to make the excerpts fill more of the screen. Concerns that this may affect the data collected using an eye tracker were allayed after receiving a confirmation from *Tobii* support that there is no major adverse impact (see Appendix D).

7.6.3.3 Data collection instruments

Questionnaire:

In the current experiment, all subjects were given the questionnaire in paper form after seating themselves at the eye tracker. The questionnaire was pre-tested on four graduates from DCU to eliminate inconsistencies in the questions and some changes were made as mentioned below. There were two sections in the questionnaire, which can be found in Appendix E.

The first section dealt with the 11 excerpts shown in Figure 7.6, and contained questions focussing on the content of the excerpts, with one open question on a dialogue subtitle and one on a CVNC for each excerpt. The questions appeared in the same order as the items appeared in the excerpt, so if, for example, the subtitle on which the question was asked appeared before the CVNC, then the question about the subtitle also appeared first in the questionnaire. This was done after the questionnaire was piloted and some subjects noted that it was confusing if the questions were not in the same order as the items that were being asked about appeared in the excerpt. The questions in this section dealt with perceived understanding of the CVNC (declared PCEs relating to CVNC), actual understanding of the CVNC (PCEs relating to CVNC), retention of dialogue subtitle content (PCEs relating to subtitle), and perception of dialogue subtitle speed (processing effort).

The questions on the CVNCs were adapted from the format of questions used in elicited perception studies outlined in the literature review (Chiaro 2004, Antonini 2005, 2007, Bucaria 2005, Bucaria and Chiaro 2007). Subjects were asked to rate their level of understanding of the CVNC on a Likert scale between one and six, with one representing that they did not understand the CVNC at all, and six representing that they understood it fully. In the design, a scale of one to six was desirable for later analysis, as the lack of a middle point meant that all answers could be easily divided into two groups, declared understood, or declared not understood. The Likert scale question was followed by an open question, asking the subject a specific question about the CVNC to determine their understanding of it.

The question on subtitle speed was adapted from Jensema (1998), and was originally designed to measure viewer reaction to different subtitle speeds. Initially it was decided to show the scale from Jensema (ibid, see Section 6.2.1.1) at the start of the questionnaire. However, comments from the pre-test of the questionnaire indicated that these explanations were “unnecessarily long” and “tiresome”, so the scale was not included in the final version of the questionnaire.

The second and final section of the questionnaire asked which strategy viewers thought should be used to translate CVNCs, as well as some questions on their opinions of the use of pop-up gloss.

Eye tracker:

The eye tracking hardware and software used in the experiment have been previously discussed (see Chapter 5). For the purposes of the experiment, scenes were defined using *ClearView* for the duration of time that each subtitle was on screen. This was considered a more consistent measure than that used in the pilot study³⁵, and it allowed for more precise comparative analyses to be made with the eye tracking results. Within each scene, two AOIs were defined for the analysis in the current experiment: subtitle and pop-up gloss.

³⁵ In the pilot study, a scene was defined as the duration the CVNC was on screen, resulting in scenes of different lengths for each excerpt.

7.6.3.4 Subjects

The sample for the present study was recruited mainly from members of the DCU Anime Society and subjects from a previous study on anime, manga and videogames (O'Hagan 2008b), with several other subjects obtained from postgraduate students of the School of Applied Language and Intercultural Studies at DCU. This approach was taken in an effort to maintain proximal similarity within the sample, as well as being dependent on the availability of resources. A recruitment email (see Appendix F) was sent to the DCU Anime Society mailing list, students who had participated in the pilot study, as well as the humanities postgraduate student mailing list. Care was taken to show potential subjects the legitimacy of the study by including a preface message by the thesis supervisor. The experiment was described as a study into the perception of TV anime. The deliberately broad description of the focus of the experiment was followed for the entirety of the experiment, avoiding reference to CVNCs or subtitles in an effort to avoid a disproportionate amount of attention being paid to either semiotic channel (see also Section 7.5.2.4). This follows the procedure used by d'Ydewalle *et al.* (1991), where subjects were told they were involved in a study of emotional reactions to AV content rather than the real topic of the study, the reading of subtitles. Prior to the experiment, an email questionnaire was sent to subjects (see Appendix G) asking for some biographical information and their opinions on anime, as well as viewing habits.

All subjects chose an answer among “yes”, “no”, “no opinion” and “don't know what it is”. Of the twenty subjects, only one (whose eye data, because of the high amount of missing data, was not included in the study anyway) said she did not like anime, while thirteen said they did, five had no opinion on it and one did not know what it was. The majority of the sample could therefore be said to have no negative feelings towards anime, and may be a representative sample of fans and occasional viewers of anime.

As highlighted previously, some physical factors can affect the results obtained with the eye tracker, particularly pupil results (Verney *et al.* 2001). On the basis

of the questionnaire used by Schultheis³⁶ (2004), subjects were asked if they had any of the following medical conditions before taking part in the experiment:

- any condition which adversely affects vision or reading
- substance abuser
- diabetes
- brain or serious head injury

In the end 20 subjects were recruited, 10 male and 10 female, and all were native speakers of English. The gender balance was an intentional part of the experiment design because of previous research findings (d'Ydewalle *et al.* 1987) that indicated differences in reading behaviour with subtitled AV content, with female subjects looking at subtitles for longer than male subjects. However, the more recent results obtained by Jensema *et al.* (2000b) showed no significant difference between male and female viewers in their reading behaviour, and for this reason, the gender divide was disregarded in subsequent calculations. There is also no gender difference displayed in terms of the pupillary response to processing effort (Beatty 1982).

The average age of subjects for the present study was 24.75 years ($SD=7.71$), which can be considered as within the representative age group of the anime viewing population. This is supported by the results of a recent online survey of 776 anime fans which revealed the majority of respondents (40%) were aged between 19 and 25, and 20% were aged between 26 and 30 (Igarashi 2008). Four subjects had a basic level of Japanese, but declared that they were fully dependent on the subtitles while viewing the excerpts. Nonetheless, this may have had an effect on the data gathered in the study.

The number of subjects in the experiment was also dictated by the practical implications of the vast volumes of data which the eye tracker generates and more particularly the lack of a prior study using pupillometry with subtitled AV content, meaning that in order to practically deal with the data volume and its analysis, a manageable number of subjects was recruited.

³⁶ Schultheis' (*ibid*) study investigated the processing effort required to read texts of varying degrees of difficulty using pupillometry.

7.6.4 Procedure

A protocol was prepared in advance of the experiment and was followed throughout each experiment session. The experiment took place in the Advanced Translation Research laboratory in DCU. An attempt was made to provide a relatively constant level of ambient light for each subject by having the same light on and the blinds shut. Individual appointments were made with each subject and they came to the laboratory individually. Upon arrival at the laboratory, the subject was asked to read the plain language statement and to sign the informed consent form. The subject was then asked to take a seat in front of the eye tracker.

After the subject was seated at the monitor, the researcher assisted in calibrating the eye tracker to his/her eyes. While Duchowski (2003) notes that it is not necessary and sometimes even undesirable that subjects in diagnostic studies are made aware of the presence of an eye tracker, the need to perform a calibration of the monitor made this unfeasible in the present study. The calibration process is outlined in Section 5.3.1. With the calibration successfully completed, the subject was given a brief description of the series, copied from the text on the back of the DVD box (see Figure 7.7).

The first excerpt shown to subjects supplemented this description, as it came from the introduction to an episode and gave a brief explanation of the background of the series. After seeing the excerpt, subjects were asked to answer questions on it, before moving onto the next excerpt. While the lack of contextualisation for each of the excerpts may be seen as a drawback in some experiment set ups, the nature of the AV content used in the current experiment made the lack of contextualisation of minor importance because of the self-contained nature of the excerpts used.

10TH GRADE CLASS + 11-YEAR-OLD TEACHER
(YOU DO THE MATH.)

Good news? She's an MIT grad. Bad news? She's a 11-year-old MIT grad. So while Becky Miyamoto may be intellectually able, this MIT prodigy is painfully ill-equipped to deal with a group of temperamental teens, especially THIS group. There's the bitchy heather, the ADD spazz, the angry nerd, the identical twins, the invisible girl, the freaky class president, the drama geek, the Nancy Drew, the gamer, the princess...whew!

Add to that her pathetically-abused stuffed bunny buddy and idiotic aliens watching her every move and it's no wonder Becky is prone to crying fits, tirades, and flipping the class the bird.

Don't miss the Grade-A comedy from the brains behind *Negima?* – *Paniponi Dash!* It's a lesson in comedy that'll just kill ya.

Figure 7.7. Text from description on back cover of *Paniponi Dash!* (2007) DVD

Each subject saw a mixture of excerpts in the treatment condition and the normal condition (see Section 7.2.1). As outlined in Table 7.6, the first ten subjects were shown each *odd-numbered* excerpt in the treatment condition, while the second ten subjects were shown each *even-numbered* excerpt in the treatment condition.

<i>Excerpt no.</i>	1	2	3	4	5	6	7	8	9	10	11
Subjects 1-10	T	N	T	N	T	N	T	N	T	N	T
Subjects 11-20	N	T	N	T	N	T	N	T	N	T	N

Table 7.6. Experiment condition in which subjects saw each excerpt ("T" for treatment, "N" for normal)

7.7 Summary

In this chapter, the methodological considerations that informed the design of the experiment were elaborated. A pilot study that was carried out to increase familiarity with the eye tracker was described. Factors that may affect the validity of the experiment were considered, and the independent and dependent variables to be used in data analysis were selected. The design and procedure followed in the experiment were then described in detail. The following chapter provides an analysis of the results of the experiment.

CHAPTER 8. **D**ATA ANALYSIS

The previous chapter dealt with the methodological considerations that informed the experiment design, and then went on to describe the experiment in detail. This chapter discusses the data from the experiment under two headings, depending on the instrument used to collect the data. The questionnaire data is discussed first, followed by the data collected with the eye tracker. The chapter is concluded with a general discussion of the data.

8.1 Questionnaire data

As noted in the previous chapter, the data from the questionnaire was used to compare the difference between subjects who watched the same excerpt in one of two experiment conditions:

1. *Treatment:*

In this condition, the excerpt contained pop-up glosses, as well as some additional captions of the background text. Using Nornes' (1999) terminology, these excerpts could be classified as using abusive subtitles. The CVNC being examined in the excerpt was explained by a pop-up gloss (see Figure 7.1, left).

2. *Normal:*

In this condition, there was no pop-up gloss at all in the excerpt, and no extra verbal reference was made to the CNVC except those contained in the dialogue subtitles (see Figure 7.1, right). These excerpts could be classified under Nornes' (1999) terminology as using corrupt subtitles.

8.1.1 Hypotheses about questionnaire data

The premise of the experiment was that subjects who saw the excerpt in the treatment condition would experience improved PCEs related to CVNCs, while PCEs were expected to be lower for subtitles. This is because it was expected that the presence of pop-up gloss, while providing extra information about CVNCs, would also distract the subjects' attention from the subtitle area. Processing effort was expected to be higher in the treatment condition, and subjects were expected to experience subtitles as being quicker when a pop-up gloss was present. In total, four hypotheses were tested using the questionnaire data.

1. As the presence of pop-up gloss was expected to provide additional information to subjects about the CVNC in each excerpt, it was hypothesised that the number of subjects experiencing PCEs in the treatment condition would be higher than those in the normal condition. This would result in subjects watching the excerpt in the treatment condition obtaining more correct answers to the CVNC questions:

Hypothesis [1]

H_0 : Mean correct answers to CVNC question_{abusive} \leq Mean correct answers to CVNC question_{corrupt}

H_1 : Mean correct answers to CVNC question_{abusive} $>$ Mean correct answers to CVNC question_{corrupt}

2. Subjects' confidence in the accuracy of the information they gathered from the excerpt was also expected to be higher in the presence of pop-up gloss. Therefore, it was not only hypothesised that the number of subjects in the treatment condition experiencing PCEs relating to CVNCs would be higher than those in the normal condition, but there would also be a higher number of subjects who had confidence that the information communicated about CVNCs in the treatment condition was accurate. This would be reflected in a higher number of subjects saying that they understood the CVNC in the treatment condition:

Hypothesis [2]

H_0 : Mean declared understanding of CVNC_{abusive} \leq Mean declared understanding of CVNC_{corrupt}

H_1 : Mean declared understanding of CVNC_{abusive} $>$ Mean declared understanding of CVNC_{corrupt}

3. While the number of subjects experiencing PCEs in the treatment condition was expected to be higher in terms of CVNCs, the use of pop-up gloss was expected to distract attention from the subtitles and result in a lower number of subjects experiencing PCEs relating to subtitles in the treatment condition. This would result in subjects who watched the excerpt in the normal condition obtaining more correct answers in the subtitle questions:

Hypothesis [3]

H_0 : Mean correct answers to subtitle question_{treatment} \leq Mean correct answers to subtitle question_{normal}

H_1 : Mean correct answers to subtitle question_{treatment} $>$ Mean correct answers to subtitle question_{normal}

4. The final hypothesis made about the questionnaire data was related to processing effort. As highlighted previously, it was expected that subjects who require more effort to watch an excerpt would experience the subtitle speed as being fast or too fast to read. As it was assumed that subjects would require more processing effort to view excerpts in the treatment condition, it was expected that a greater number of subjects would say the subtitles were fast in the treatment condition:

Hypothesis [4]

H_0 : Mean declared fast_{abusive} \leq Mean declared fast_{corrupt}

H_1 : Mean declared fast_{abusive} $>$ Mean declared fast_{corrupt}

In the following, the results of the questionnaire are discussed in the same order as the hypotheses listed above.

8.1.2 PCEs relating to CVNCs

The data in Figure 8.1 shows the number of subjects that actually understood each CVNC, depending on the experiment condition. As the graph illustrates, the number of correct answers in the treatment condition was higher than the normal condition in all instances except for Excerpt 7. A paired t-test was carried out to compare the number of correct answers in each condition, the result of which was statistically significant ($t(10)=3.297$, $p<0.01$). As expected, the mean number of correct answers in the treatment condition ($M=4.91$, $SD=2.77$) was higher than in the normal condition ($M=3.18$, $SD=2.40$).

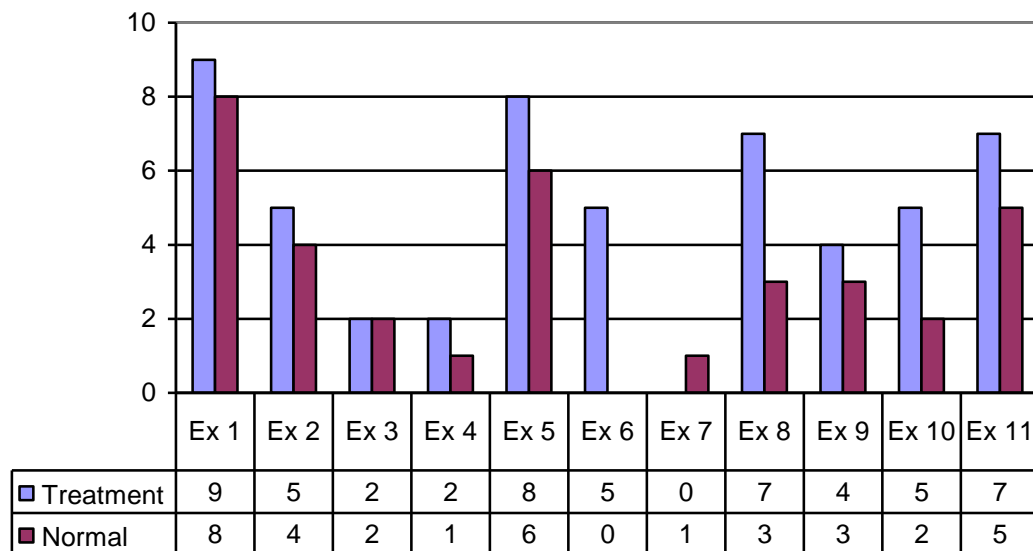


Figure 8.1. Number of correct answers to CVNC question for each excerpt according to experiment condition

Only three CVNCs were correctly interpreted by more than half of all subjects (i.e. from both treatment and normal conditions), those found in Excerpt 1, Excerpt 5 and Excerpt 11. As explained in the excerpt selection process outlined previously, the CVNCs in Excerpts 1 and 11 were both based on Western culture, so the higher number of correct answers was expected in those cases.

Excerpt 1 was taken from the beginning of an episode of *Paniponi Dash!* (2007). In the excerpt, a student explains that she is going to expose Rebecca Miyamoto's secret. The type of language used in this excerpt, both visual verbal (in the form of the subtitles, see highlighted section in Table 8.1) and visual nonverbal (the use of a gun sight camera angle, see Figure 8.2), link it to the espionage genre, and more particularly to the introduction sequence used in the *James Bond* film series.

<i>Spoken dialogue</i>	<i>Subtitle</i>	<i>Pop-up gloss</i>
一年 B 組綿貫響	In the name of the Intelligence Department of the Student Council,	This is an obvious spoof of the opening sequences to the James Bond movies.
生徒会諜報部の名に かけてあなたの 秘密を暴かせる	I, Hibiki Watanuki of Class 1-B, Shall reveal your secret!	

Table 8.1. Subtitles from Excerpt 1

The CVNC in this excerpt could be called a visual reference, as it makes a reference to the visuals in another piece of AV content, and it was the CVNC recognised correctly by most subjects. This may be due to the presence of clues in the verbal and nonverbal visual channel in both normal and treatment conditions, as well as the widespread popularity of the *James Bond* film series.



Figure 8.2. Screen captures from *Paniponi Dash!* (2007, left) and *The Living Daylights* (1987, right)

Even though the other Western CVNC in Excerpt 11 was correctly recognised by more than half of the subjects, there were fewer correct answers to the CVNC question than Excerpt 1, despite the popularity of the *Batman* film franchise to which the excerpt refers. In Excerpt 11, there are several cuts between a scene of Rebecca experimenting on her rabbit in a classroom, and an outdoor scene with a girl lying under a tree surrounded by cats (see Figure 7.6, number 11). There is a parallel drawn between both scenes by the subtitle (see Table 8.2). While Rebecca experiments on the rabbit, a fairy is born out of its head, and the scene with the cats is a reference to the scene in the film *Batman Returns* (1992) where Catwoman is reborn after being murdered by her employer. Unlike Excerpt 1, the dialogue in Excerpt 11 is not thematically linked to Catwoman or the film genre from which the visual quote comes, relying

entirely on the visual nonverbal channel to transfer the message. In the treatment condition, the accompanying pop-up gloss explained the scene as displayed in Table 8.2.

<i>Spoken dialogue</i>	<i>Subtitle</i>	<i>Pop-up gloss</i>
なんか生まれてるし	Now something was born.	A spoof of the movie <i>Batman Returns</i> . This is the scene where Catwoman is born.

Table 8.2. Subtitle from Excerpt 11

The pop-up gloss appears to have caused confusion for Subject 2, who answered “catgirl returns” to the question asking which character the scene made reference to. Another subject from the same condition appears to have misread the question, answering “Batman Returns (spoof)” (Subject 8). The same answer (“Batman Returns”) was also given by Subject 11 who watched the excerpt in the normal condition, despite the fact that there was no verbal reference to the film in the normal condition.

The only other excerpt where the CVNC question was answered correctly by more than half of all subjects was Excerpt 5. The CVNC in question here was a photo frame wrapped in a black ribbon (see Figure 7.6, number 5). Similar to Excerpt 1, a plausible reason for the high number of correct answers to this question is the presence of a reference to it in the subtitle (see Table 8.3), allowing subjects to associate the reference to death with the black ribbon. It could also be a result of the widespread use of the colour black as a colour of mourning.

<i>Spoken dialogue</i>	<i>Subtitle</i>	<i>Pop-up gloss</i>
いきなり殺すな	Don't kill her off already!	In Japan, the photo frame for a deceased person's picture is decorated with a black ribbon
黒いものを取れ、 黒いの	Take that black thing off!	

Table 8.3. Subtitles from Excerpt 5

In contrast to Excerpts 1, 5 and 11, the excerpt with the lowest number of correct answers to the CVNC question was Excerpt 7, with only one correct answer from all 20 subjects. Excerpt 7 presented subjects with a visual

reference to a collection of woodcuts by the internationally renowned artist Hokusai Katsushita, 富嶽三十六景 (*fugaku sanjūrokkei*, Thirty six views of Mount Fuji) (see Figure 7.6, number 7). The only subject who answered this question correctly watched the excerpt in the normal condition, and thus was given no extra information about the CVNC except the visual reference in the visual nonverbal channel. Excerpt 7 is the only instance where more subjects who viewed the normal condition of the excerpt answered the CVNC question correctly, although two subjects who viewed the treatment condition of the excerpt did comment that they were familiar with the work of the artist, but couldn't remember his name.

Comparing the difference between the numbers of correct answers in both experiment conditions, the excerpt displaying the largest gap is Excerpt 6. The CVNC question here asked subjects the purpose for which *karakuri* dolls were originally designed. *Karakuri* dolls (絡繰人形, *karakuri ningyō*) are dolls that move via mechanisms made from string and springs. The one in the scene is described by the pop-up gloss as “the classic form designed during the Edo period to serve tea”. Although the doll can be seen carrying a bowl of green tea in the excerpt (see Figure 7.6, number 6), this may not have been easily recognisable as tea to the subjects.

8.1.3 Declared PCEs regarding CVNCs

Declared PCEs were measured using a subjective Likert scale question about the same CVNC in each excerpt on which the open question was based. Figure 8.3 shows the number of subjects who declared understanding of the CVNC from each excerpt in each experiment condition. Subjects who rated their understanding between one and three were counted as “declared not understood”, while those who rated their understanding between four and six were counted as “declared understood”. The results of the t-test suggest a significant effect for experiment condition ($t(10)=2.39$, $p<0.05$), with the treatment condition corresponding with a higher number of subjects declaring understanding of CVNCs ($M=4.09$, $SD=2.26$) than the normal condition ($M=2.64$, $SD=2.5$).

The only exception is Excerpt 1, where more subjects from the normal condition declared understanding of the CVNC than the treatment condition. Once again, the popularity of the *James Bond* film series probably influenced this result. For the questions on declared PCEs, only two excerpts were declared understood by more than 50% of all subjects, Excerpt 1 and Excerpt 11, both Western pop-culture references. The fact that these references were based on Western pop-culture probably meant that subjects were more familiar with them; increasing their confidence in the accuracy of the information they obtained from the excerpts.

There were only two excerpts where the difference in declared understanding between conditions exceeded three subjects: Excerpt 2 and Excerpt 5. In both cases, more subjects in the treatment condition declared their understanding of the CVNC.

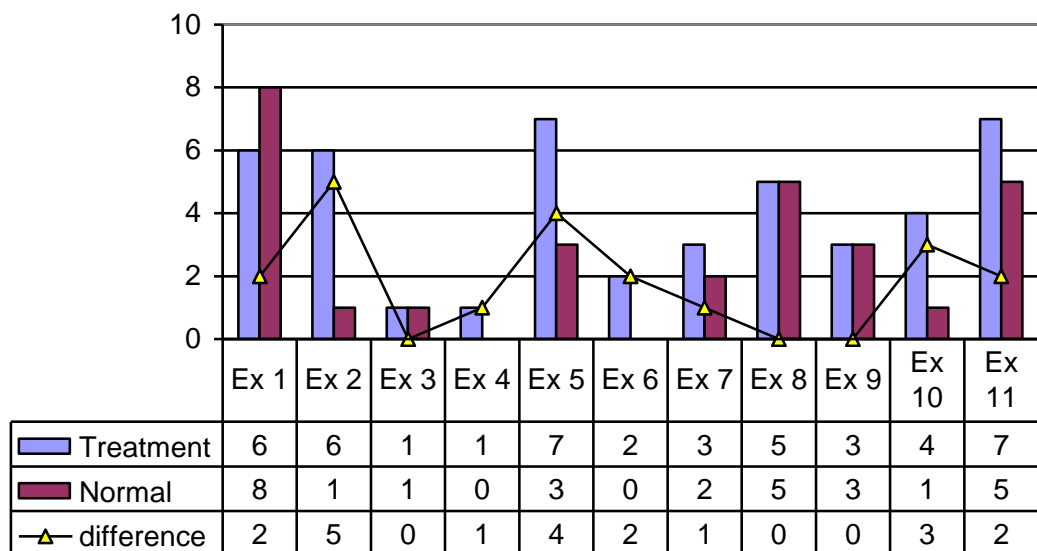


Figure 8.3. Declared understanding of CVNC in each excerpt according to experiment condition and difference between conditions

The CVNC in Excerpt 2 was a 2-chan emoticon depicting shock (see Figure 7.6, number 2). In the excerpt Rebecca's face changes into an emoticon after her students catch her out for losing something. The emoticon is one of many that are in common use in Japan, and it depicts an open-mouthed face. Despite the high discrepancy in declared understanding for this item, Figure 8.4 illustrates that the number of correct answers to the open CVNC question differed by only

one subject. The discrepancy between declared and actual understanding in the normal condition may be because of subjects' lack of familiarity with this form of emoticon, and the highly specific nature of emoticons that enables accurate interpretation without the viewer necessarily needing to be familiar with them.

8.1.3.1 Relationship between declared and actual PCEs related to CVNCs

As a final part of the analysis of PCEs, the correlation between perceived and actual PCEs was calculated. This correlation can be used as an indicator of how close the subjects' *impression* of the accuracy of information garnered from the excerpt was to the *actual* accuracy of the information. Both scores were strongly correlated in both the treatment (.78) and normal (.81) conditions, indicating that subjects' average impression of the PCEs they experienced tended to be accurate.

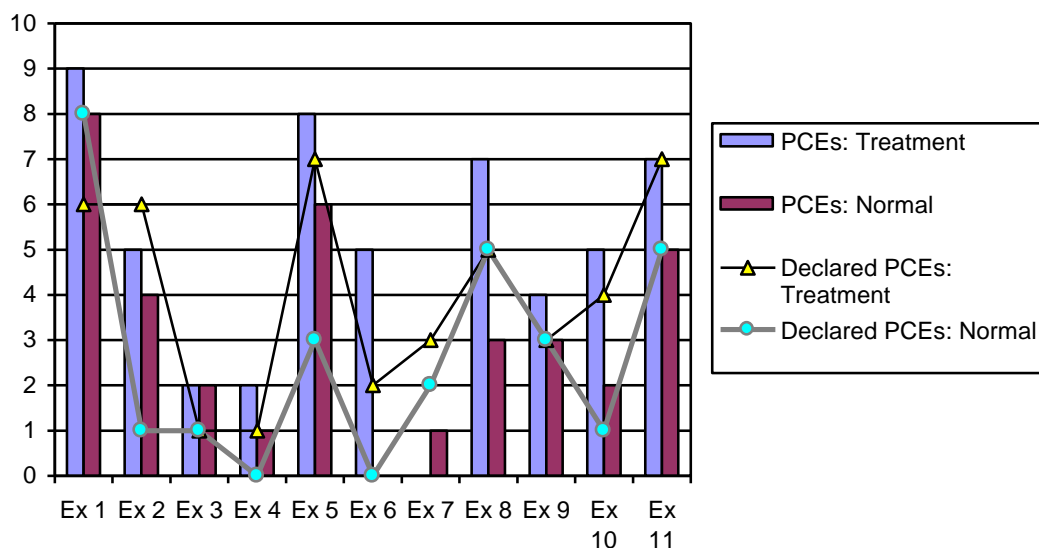


Figure 8.4. Actual and declared PCEs related to CVNCs for each excerpt according to experiment condition

In most cases, declared PCEs were lower than actual PCEs, standing in contrast to Antonini's (2005, 2007) and Bucaria's (2005) findings, where declared understanding was consistently higher than actual understanding. In fact, the only excerpt where declared PCEs were higher than actual PCEs in

both experiment conditions was Excerpt 7 (see Figure 8.4), which as noted above, was a visual reference to a painting that several subjects said they recognised although they could not remember the name of the painter. Subjects declaring that they understood the CVNC, yet not being able to name the painter may have been the cause of the discrepancy between declared and actual PCEs relating to the CVNC in this excerpt.

8.1.4 PCEs relating to subtitle

The data in Figure 8.5 displays the number of correct answers to the subtitle question for each excerpt and each experiment condition, as well as the difference between the numbers of correct answers in each condition. While the average number of correct answers in the normal condition ($M=7.55$, $SD=2.16$) was higher than in the treatment condition ($M=6.64$, $SD=2.54$), the difference in means was not statistically significant ($t(10)=0.959$, $p>0.05$).

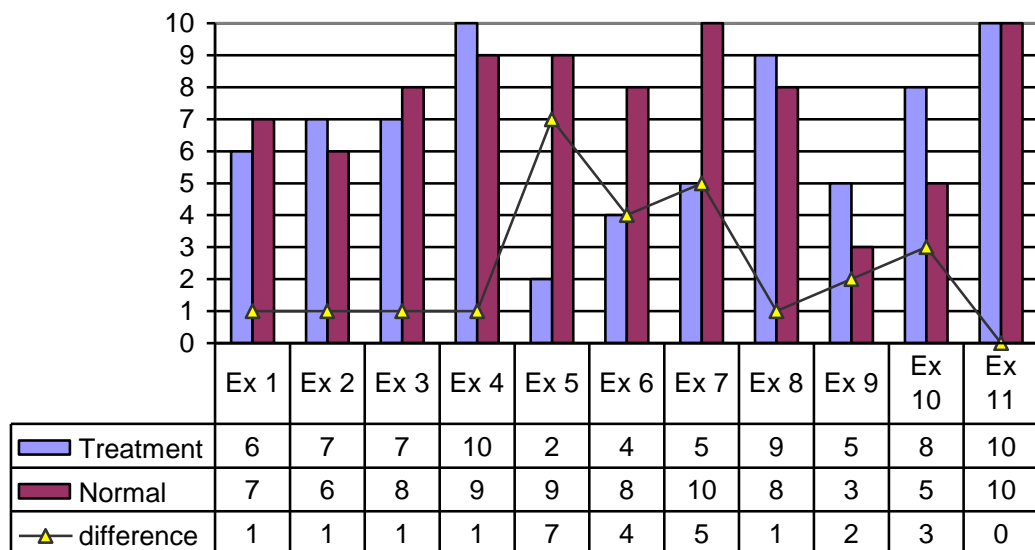


Figure 8.5. Number of correct answers to subtitle question for each excerpt according to experiment condition and difference between conditions

Although the lack of significant difference between the means of each condition was not expected, on closer inspection three excerpts become conspicuous because of the larger difference between conditions (i.e. a difference of four or more subjects), Excerpt 5, Excerpt 6 and Excerpt 7. Two distinguishing features that may have led to the higher differences between conditions in these instances were discerned.



Figure 8.6. Screen capture of subtitles from Excerpt 5 (top) and Excerpt 7 (bottom) on which subtitle content question was based (*Paniponi Dash! 2007*)

In Excerpt 5 and Excerpt 7 the subtitles that were asked about in the subtitle content questions were on screen at the same time as a pop-up gloss (see Figure 8.6). Along with Excerpt 1, these were the only excerpts where the subtitle that was the topic of a question was on screen at the same time as a pop-up gloss. The fact that the subtitles were two lines long may have meant that there was not enough time for some subjects to make a proper reading of both subtitle and gloss or to commit the information to memory. However, for Excerpt 1, where the subtitle was also two lines (see Table 8.1), the difference in correct answers between conditions is very low.

There are two possible reasons for this. One is the relative length and complexity of the pop glosses involved; in Excerpt 1, only one three-line pop-up gloss is displayed on screen, while in Excerpt 5 there are two glosses of three lines each on screen (Figure 8.6, top), and Excerpt 7 has unusual semantic

blocking³⁷ with the use of a hyphen separating the word ‘sarcastic’ between two lines (Figure 8.6, bottom), possibly demanding more processing effort be allocated to the gloss at the expense of the subtitle. Another possible reason is the visual context of the scenes where the subtitles are present. While the scene in Excerpt 1 has relatively simple visual activity on the screen, with there usually being only one central point of focus at any time, in Excerpt 5 and Excerpt 7, the scenes are more visually complex, plausibly demanding more division of attention, although there was little difference in the percentage gaze time in the subtitle area in all three excerpts for the treatment condition (41%, 42% and 54% respectively). This could be investigated further by carrying out a more controlled experiment in the future.

The reason for the discrepancy in Excerpt 6 appears to be related to subjects mistaking the identity of characters in the excerpt. The subtitle content question for this excerpt asked why the rabbit character thought drinks from the vending machine were warm. The reason the rabbit gives is “perhaps they just put them in, so they’re not chilled yet” (see Figure 8.7, top). However, several respondents incorrectly answered with the reason given by the cat character in the same scene, “body heat” (see Figure 8.7, bottom). This discrepancy does not appear to have had anything to do with the presence of pop-up gloss, as there is no pop-up gloss on screen at the same time as either subtitle, or any of the subtitles immediately around them.

Instead it seems that subjects may have mistaken what each character said, as both rabbit and cat characters are on screen at the same time as the subtitle referring to body heat. Although it should be noted that the subtitle referring to body heat ends with a “meow”, clearly marking the cat as its speaker.

³⁷ Semantic blocking refers to the way in which the on-screen text is segmented. For ease of reading, line breaks should generally be made at the highest possible syntactic nodes (Karamitroglou 1998).



Figure 8.7. Screen captures from Excerpt 6 showing rabbit and cat characters

8.1.5 Processing effort

Processing effort was measured using a question that asked subjects to give their subjective judgement of subtitle speed. This method was used to avoid asking subjects directly about the effort they were experiencing. Furthermore, it was assumed that if subtitles were perceived as being too quick, the subjects would have required more cognitive effort to process the excerpt. Subjects who rated the subtitle speed between four and five were categorised as having declared the subtitle speed as “fast”.

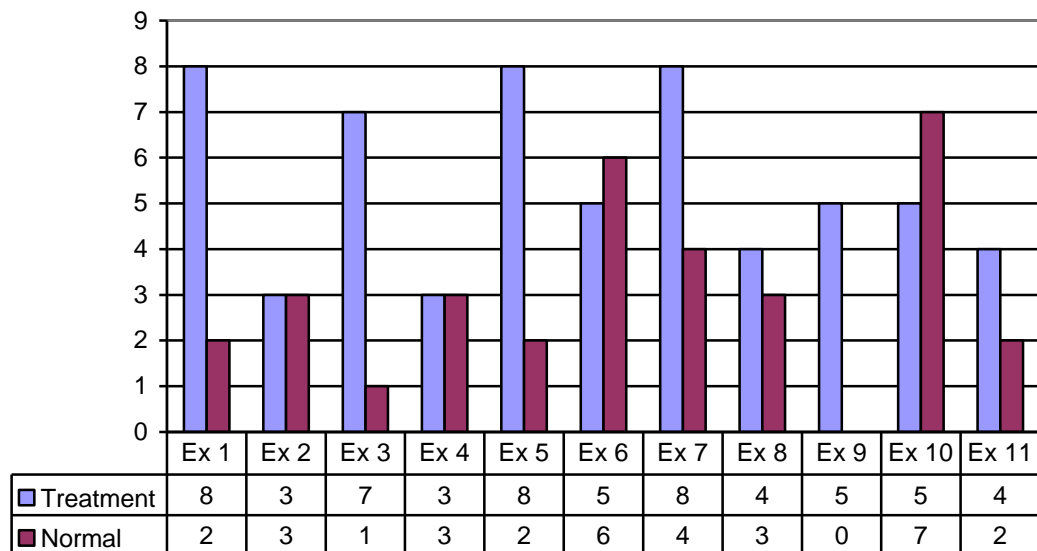


Figure 8.8. Number of subjects who declared subtitle speed "fast" for each excerpt according to experiment condition

The graph in Figure 8.8 shows the number of subjects who declared the subtitle speed as "fast" from each condition and excerpt. A significant effect was found for the experiment condition ($t(10)=2.673$, $p=0.02$), with a mean of 5.45 "fast" answers for the treatment condition ($SD=1.97$), and the lower mean of 3 for the normal condition ($SD=2.05$). This indicates that the presence of pop-up gloss in the excerpts made subjects perceive the subtitles as faster, indicating that more effort was required to process the excerpts when pop-up gloss was used.

Nevertheless, a factor that should not be ignored here is that the difference between conditions depends heavily on the excerpt numbers. Whereas all odd numbered excerpts (which were viewed in the treatment condition by subjects 1-10) show a large difference between the means in the treatment ($M=6.67$, $SD=1.75$) and normal ($M=1.83$, $SD=1.33$) conditions, this difference decreases drastically for the even numbered excerpts (which were viewed in the "treatment" condition by subjects 11-20), with the number of "fast" answers appearing in the "normal" condition ($M=4.4$, $SD=1.95$) nearly equalling those in the "treatment" condition ($M=4$, $SD=1$). This places the reliability of the t-test carried out above into question, as it seems that more of the subjects 1-10 rated subtitles as fast independent of the condition, which may have resulted in an unwanted bias affecting the figures.

In an effort to ascertain the reason for subjects 1-10 generally rating subtitle speeds as quicker than subjects 11-20, the answers to the questions in the preliminary email questionnaire (see Appendix G) relating to AVT preferences and viewing habits were considered. While the number of subjects who watched subtitled AV content once a month or more was almost equal with 7 from subjects 1-10 and 8 from subjects 11-20, subjects 11-20 showed a higher preference for subtitling as their favoured mode of AVT. 7 subjects from subjects 11-20 chose subtitling as their preferred form of AVT versus only 4 subjects from subjects 1-10. This preference for subtitles may mean that subjects 11-20 are more accustomed or comfortable with viewing subtitled AV content, thus resulting in a lower level of processing effort and perceived speed of subtitles.

8.1.6 Summary of questionnaire analysis

The analysis of the questionnaire data can be summarised as follows:

- The mean number of subjects who experienced PCEs relating to CVNCs was significantly higher when pop-up gloss was used in an excerpt.
- The mean number of subjects who perceived experiencing PCEs relating to CVNCs was also significantly higher when pop-up gloss was used in an excerpt.
- There was a positive correlation between actual and perceived PCEs relating to the CVNC, and declared PCEs were generally lower than actual PCEs.
- Although the mean number of subjects who experienced PCEs relating to subtitles was lower when pop-up gloss was used in an excerpt, this difference was not statistically significant. As such hypothesis [3] could not be proved. However, the data suggested that instances where there was a more pronounced difference in the PCEs experienced relating to subtitles may have been caused by the presence of pop-up gloss on screen at the same time as the subtitle on which the question was asked, or confusion related to characters in the excerpts.
- A significantly greater number of subjects rated the subtitle speed fast when pop-up gloss was used in the excerpt, suggesting that more processing effort was required to view excerpts with pop-up gloss.

However, subjects 1-10 showed a general tendency to rate subtitles as faster than their counterparts.

8.2 Fixation-based and pupillometric data

As explained previously, fixation-based and pupillometric data were used to make within-subject comparisons, so instead of comparing the difference between the same excerpt in different conditions, in this section of the data analysis a comparison is made between the differences in four experiment conditions for each subject. The number of subtitles in each condition seen by subjects 1-10 differed from those seen by subjects 11-20, because of the different forms of the excerpt they saw. Accordingly analyses were made on different subtitle sets for different subjects. Table 8.4 shows the number of subtitles in each condition seen by the subjects.

<i>Condition</i>	<i>Subjects 1-10</i>	<i>Subjects 11-20</i>
Solo/one line	147	162
Solo/two line	23	29
Gloss/one line	24	9
Gloss/two line	8	2
<i>Total</i>	202	202

Table 8.4. Number of subtitles in each condition for subjects 1-10 and 11-20 respectively

Section 7.6.3.4 outlined reasons that the eye data from some subjects was not included in the fixation-based or pupillometric analysis. Taking that into account, the total number of valid data entries for each form of analysis and experiment condition is displayed in Table 8.5.

<i>Condition</i>	<i>Fixation-based analysis</i>	<i>Pupillometric analysis</i>
Solo/one line	2163	2016
Solo/two line	364	341
Gloss/one line	231	207
Gloss/two line	70	62

Table 8.5. Total number of data entries used in fixation-based and pupillometric analyses according to experiment condition

8.2.1 Hypotheses about fixation-based and pupillometric data

Fixation-based and pupillometric data was used to determine the processing effort experienced by subjects while watching excerpts, and several comparisons and analyses were made. It was expected that the use of pop-up gloss and two-line subtitles would necessitate a higher level of processing effort than one-line subtitles with no pop-up gloss. Processing effort was expected to be at the highest level when a two-line subtitle was on-screen with pop-up gloss. The presence of pop-up gloss was also expected to decrease the amount of processing time allocated to the subtitles. To investigate these assumptions, several hypotheses were tested.

5. The first set of comparisons was made using fixation-based measurements. The *percentage of skipped subtitles* was expected to be higher when there was a pop-up gloss on screen. The percentage of skipped subtitles was also expected to be significantly higher for one-line than two-line subtitles, as per the results in d'Ydewalle and de Bruycker (2007).

Hypothesis [5a]

H_0 : % of skipped subtitles_{gloss} \leq % of skipped subtitles_{solo}

H_1 : % of skipped subtitles_{gloss} $>$ % of skipped subtitles_{solo}

Hypothesis [5b]

H_0 : % of skipped subtitles_{one-line} \leq % of skipped subtitles_{two-line}

H_1 : % of skipped subtitles_{one-line} $>$ % of skipped subtitles_{two-line}

6. As the processing effort needed to be distributed between pop-up gloss and subtitles, the *percentage gaze time* in the subtitle area was expected to be lower when there was a pop-up gloss on screen. Again d'Ydewalle and de Bruycker's (ibid) analysis of one vs. two-line subtitles was used for comparison here, as they found that significantly less time was spent in one line than two-line subtitles.

Hypothesis [6a]

H_0 : % gaze time in subtitles_{solo} \leq % gaze time in subtitles_{gloss}

H_1 : % gaze time in subtitles_{solo} $>$ % gaze time in subtitles_{gloss}

Hypothesis [6b]

H_0 : % gaze time in subtitles_{two-line} \leq % gaze time in subtitles_{one-line}

H_1 : % gaze time in subtitles_{two-line} $>$ % gaze time in subtitles_{one-line}

7. *Mean fixation duration* was expected to be lower in instances where there was a pop-up gloss on screen, as the extra information provided in the pop-up gloss was expected to cause shorter fixations. Also, fixations in one-line subtitles were expected to be significantly shorter than those in two-line subtitles, following on from the results in d'Ydewalle and de Bruycker (2007).

Hypothesis [7a]

H_0 : mean fixation duration_{solo} \leq mean fixation duration_{gloss}

H_1 : mean fixation duration_{solo} $>$ mean fixation duration_{gloss}

Hypothesis [7b]

H_0 : mean fixation duration_{two-line} \leq mean fixation duration_{one-line}

H_1 : mean fixation duration_{two-line} $>$ mean fixation duration_{one-line}

8. The final fixation-based measurement that was used is the *word fixation probability*. D'Ydewalle and de Bruycker (ibid) did not find any significant difference between one and two-line subtitles using this measurement, and it was expected that the result would be repeated here, while it was expected that the presence of pop-up gloss would result in a significantly lower word fixation probability than instances where there was no pop-up gloss present.

Hypothesis [8a]

H_0 : word fixation probability_{solo} \leq word fixation probability_{gloss}

H_1 : word fixation probability_{solo} $>$ word fixation probability_{gloss}

Hypothesis [8b]

H_0 : word fixation probability_{two-line} \neq word fixation probability_{one-line}

H_1 : word fixation probability_{two-line} \approx word fixation probability_{one-line}

9. In terms of pupillometric measurements, the following results were expected. In both one line and two-line subtitles, the *median pupil size* of subjects was expected to be significantly larger when there was pop-up gloss present on screen than when there was no pop-up gloss on screen. Although there have been no published results using pupillometric measurements with subtitled AV content, it was expected that the increased processing effort induced by two-line subtitles, which has been indicated by other, fixation-based measurements (d'Ydewalle and de Bruycker 2007), would be reflected in the results, with median pupil size being larger for two-line subtitles.

Hypothesis [9a]

H_0 : median pupil size_{gloss} \leq median pupil size_{solo}

H_1 : median pupil size_{gloss} $>$ median pupil size_{solo}

Hypothesis [9b]

H_0 : median pupil size_{two-line} \leq median pupil size_{one-line}

H_1 : median pupil size_{two-line} $>$ median pupil size_{one-line}

8.2.2 Percentage of skipped subtitles

Hypotheses [5a] and [5b] state the expectation that the percentage of subtitles skipped would be higher in the condition where a pop-up gloss was present on screen, and also when a one-line subtitle was on screen.

When no pop-up gloss was present, an average of 2.91% of subtitles were skipped, while when there was a pop-up gloss present, this figure increased to 16.96%, revealing a significant effect of the factor subtitle type, $F(1,13) = 79.7$ $p < 0.001$ $SS = 2764.57$ $MSe = 34.70$. The subtitle length was also shown to have a significant main effect, $F(1,13) = 26.8$ $p < 0.001$ $SS = 3192.71$ $MSe = 119.27$, with an average of 7.48% of one-line subtitles being skipped, while on average, only 2.39% of two-line subtitles were skipped. D'Ydewalle and Bruycker (ibid) reported similar findings for the factor subtitle length and linked the greater number of skipped one-line subtitles to the semiotic redundancy that is often found in one-line subtitles when compared with more complex two-line subtitles.

The interaction between both factors was also found to be statistically significant, $F(1,13) = 33.5$ $p < 0.001$ $SS = 1371.77$ $MSe = 40.96$, indicating that the effect of the presence of pop-up gloss was greater in the one-line condition than the two-line condition. The highest number of subtitles was skipped when a one-line subtitle was on screen with a pop-up gloss ($M = 29.46\%$), as Figure 8.9 illustrates. This could be because of the greater amount of information in the pop-up gloss relative to one-line subtitle, particularly when we consider d'Ydewalle and de Bruycker's (2007) aforementioned considerations about the semiotic redundancy of one-line subtitles.

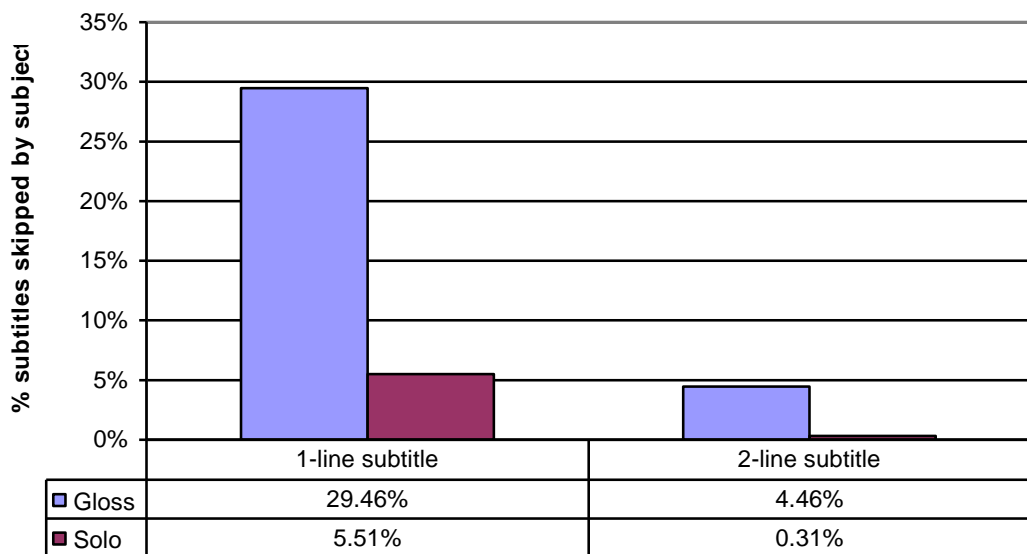


Figure 8.9. Mean percentage of skipped subtitles according to experiment condition

8.2.3 Percentage gaze time in subtitles

The percentage gaze time was expected to be higher in subtitles when there was no pop-up gloss present [6a] and when there was a two-line subtitle on screen [6b].

The novel factor under investigation in this experiment, the presence of pop-up gloss, was found to have exerted a significant effect on the percentage time spent in the subtitle area, with a smaller proportion of time spent in the subtitle area in the gloss condition ($M = 31.46\%$) than the solo condition ($M = 49\%$), $F(1,13) = 89.3$ $p < 0.001$ $SS = 3173.47$ $MSe = 48.25$. Once again following the trend in d'Ydewalle and de Bruycker (2007), subtitle length was found to have

had a significant effect on this variable, $F(1,13) = 88.5$ $p < 0.001$ $SS = 3173.47$ $MSe = 35.87$, with subjects spending on average 32.71% of their time in one-line subtitles, and 47.76% of their time in two-line subtitles.

The decrease on the percentage of time spent in the subtitle area when there was a pop-up gloss on screen may be put down to the complex structure and unusual semantic blocking in most of the pop-up glosses (e.g. Figure 8.6, bottom), which would require that some of the processing effort usually allocated to the subtitles instead be allocated to the pop-up gloss.

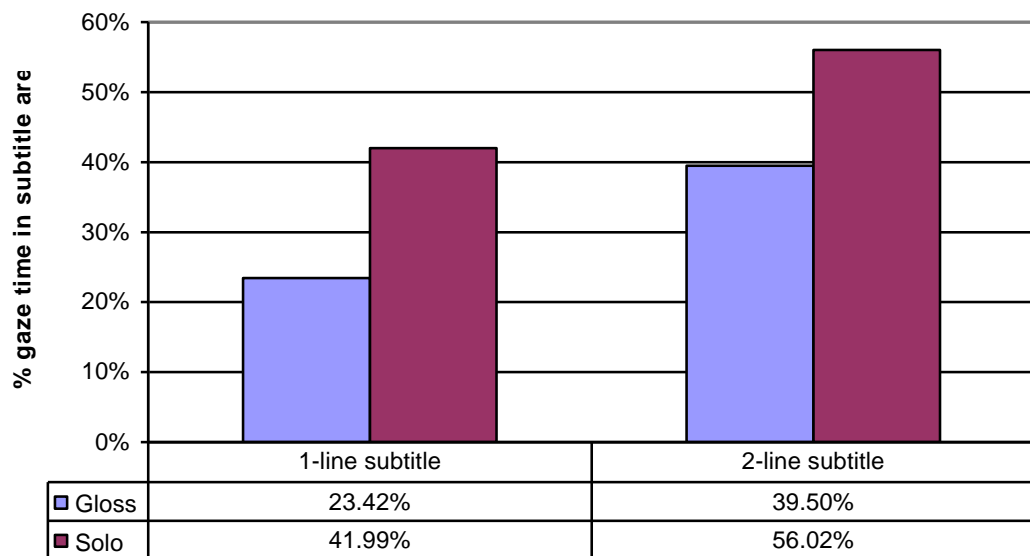


Figure 8.10. Mean percentage gaze time in subtitle area according to experiment condition

There was no significant effect found for the interaction of the subtitle type and subtitle length, $F(1, 13) = 0.682$ $p > 0.05$ $SS = 14.52$ $MSe = 23.11$. Figure 8.10 illustrates that the difference in the proportion of time spent in a one-line subtitle without a gloss is only slightly different to that spent in a two-line subtitle in the gloss condition ($M = 41.99\%$ vs. 39.5%), indicating that the presence of gloss coincides with a decreased amount of processing effort being allocated to the two-line subtitle.

8.2.4 Mean fixation duration in subtitles

As previously mentioned, the mean fixation duration was only calculated using subtitles that were actually fixated upon at least once. It was expected that mean fixation duration would be shorter when a pop-up gloss was on screen

[7a] and in one-line subtitles [7b]. Figure 8.11 shows the mean fixation duration for each of the experiment conditions.

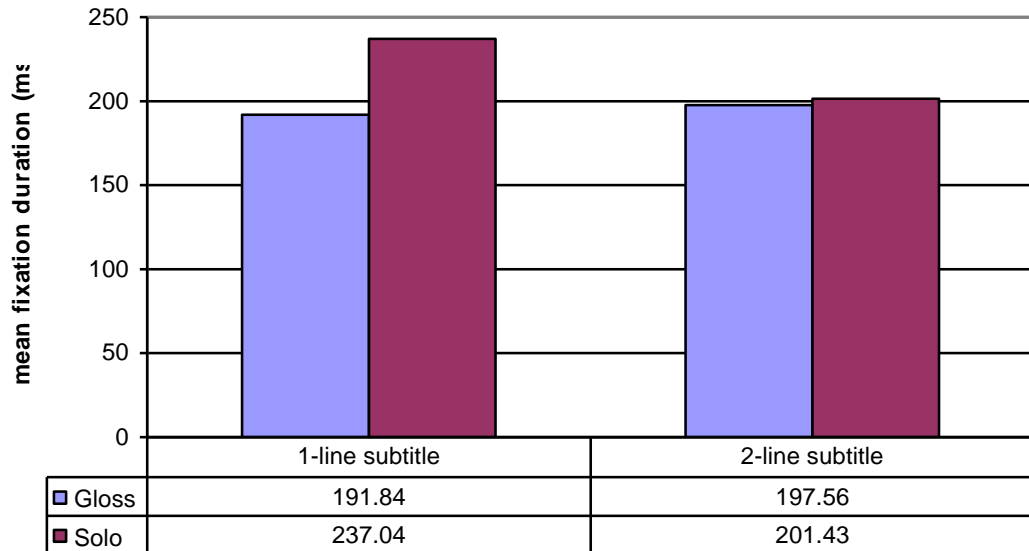


Figure 8.11. Mean fixation duration (ms) in subtitle area according to experiment condition

The analysis of fixation duration showed that there were significant main effects for the number of lines in the subtitle ($F(1,13) = 10.2$ $p < 0.01$ $SS = 3128.17$ $MSe = 306.52$), as well as the presence of pop-up gloss ($F(1,13) = 28.9$ $p < 0.001$ $SS = 8425.08$ $MSe = 291.32$). Contrary to the findings in d'Ydewalle and de Bruycker (2007), where the fixation duration in standard subtitling did not significantly differ between one and two-line subtitles, in instances where there was no pop-up gloss present, fixations in one-line subtitles were significantly longer than in two-line subtitles ($M = 237.04$ vs. 201.43 ms), $t(13) = 4.43$ $p < 0.001$. When there was a pop-up gloss on screen, the mean fixation duration was lower³⁸ in one-line subtitles (191 ms) than two-line subtitles (198 ms), the reverse of the trend when there was no pop-up gloss present. There was a significant interaction between the factors subtitle length and subtitle type ($F(1,13) = 16.4$ $p < 0.001$ $SS = 5977.66$ $MSe = 365.51$), indicating that the effect of the presence of pop-up gloss was greater in one-line than two-line subtitles. Moreover, while the fixation duration in two-line subtitles was about the same in

³⁸ Although the difference was not statistically significant, $t(13) = 1.02$ $p > 0.05$

gloss and solo conditions ($M= 198$ vs. 201 ms), in one-line subtitles, mean fixation duration was significantly longer in the solo condition than when there was a pop-up gloss on screen ($M=237$ vs. 192 ms), $t(13)=5.91$ $p<0.001$.

The overall mean fixation duration from the experiment raised some questions. Other eye-tracking studies with subtitles have found that average fixation durations tend to be shorter than those found in reading, with d'Ydewalle & de Bruycker (2007) reporting a mean fixation duration of 178 ms per word and d'Ydewalle *et al.* (1985) reporting a duration of 124 ms per word. These fixation durations are considerably shorter than the mean fixation duration for normal silent reading of English (i.e. text only) reported in Rayner (1998) of 225 ms.

The mean fixation duration for subtitles in the current experiment was 207 ms³⁹, which is closer to Rayner's (ibid) figure for normal silent reading than the figures provided above for subtitle viewing. The data in Table 8.6 illustrates that the more visually oriented tasks show higher mean fixation durations than the silent reading of text. This may provide an explanation for both the longer average fixation durations in our study, and also for the interactive effect between subtitle length and the presence of pop-up gloss. When a one-line subtitle is on screen, the least amount of text is on-screen, so the viewer spends the majority of time processing the other aspects of the excerpt. However, with the introduction of pop-up gloss, the processing of the image becomes more text-based, with the mean fixation duration decreasing and moving towards the reading end of Rayner's (ibid) scale. In this way, the presence of text-rich pop-up gloss may have shifted subjects' impressions of semiotic redundancy in the text, with the consequence of fixation behaviour following patterns found in text types where the visual verbal semiotic channel is the primary focus.

<i>Task</i>	<i>Mean fixation duration (ms)</i>
Silent reading	225
Oral reading	275
Visual search	275
Scene perception	330

Table 8.6. Approximate mean fixation duration in several tasks (adapted from Rayner ibid p373)

³⁹ This figure was calculated by calculating the mean duration for subtitles in all four conditions.

The subjects' shifting of their focus onto the visual verbal semiotic channel also provides a plausible reason for the similarities between the mean fixation duration in two-line subtitles independent of the presence of pop-up gloss. As a two-line subtitle already induces more reading-like behaviour (d'Ydewalle and de Bruycker 2007), the introduction of pop-up gloss may have less of an effect on the subjects' processing strategy for the scene.

8.2.5 Word fixation probability

Hypotheses [8a] and [8b] dealt with the probability that a word in a subtitle would be fixated upon at least once. It was expected that there would be no significant difference between one and two-line subtitles, but that the presence of pop-up gloss would result in a significantly lower probability of each word in the subtitle being fixated upon. Figure 8.12 displays the word fixation probability for each of the experiment conditions.

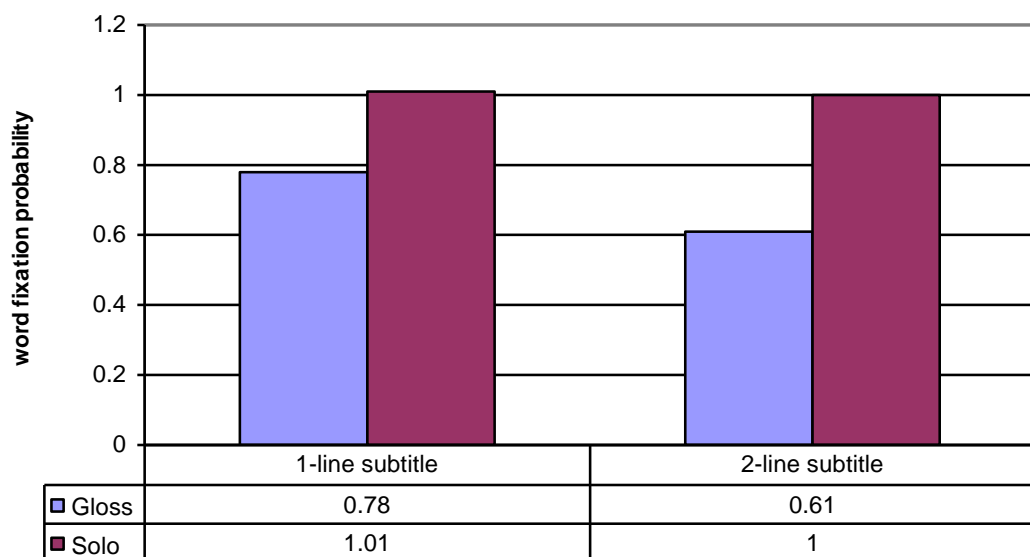


Figure 8.12. Mean word fixation probability according to experiment condition

The analysis of word fixation probability showed a significant main effect of subtitle type ($F(1,13) = 52.2$ $p < 0.001$ $SS = 1.40$ $MSe = 0.03$), with more fixations on subtitles when there was no gloss on screen than when there was a gloss on screen ($M = 1.01$ vs. 0.7). Subtitle length was also shown to have had a significant effect ($F(1,13) = 19.9$ $p < 0.001$ $SS = 0.11$ $MSe = 0.01$), with more fixations on one-line subtitles than two-line subtitles ($M = 0.9$ vs. 0.81). However,

a comparison of one and two-line subtitles in the solo condition echoes the results of d'Ydewalle and de Bruycker (2007); with the differences not being statistically significant ($M=1.01$ vs. 1), $t(13)=1.50$ $p>0.05$.

There was however a significant interaction between factors ($F(1,13) = 14.7$ $p<0.01$ $SS=0.09$ $MSe=0.01$), indicating that the effect of pop-up gloss was greater when there was a two-line subtitle on screen than a one line subtitle on screen. This resulted in the word fixation probability being lower in two-line subtitles with a gloss present than any other condition. The timing rules for subtitles tend to take subtitle length into account when deciding how long a subtitle remains on screen, so two-line subtitles tend to stay on screen for longer, allowing for more fixations the longer a subtitle is on screen. As the subtitle speed is not altered depending on the presence of pop-up gloss, when these appear on screen, there is less time to fixate and process the subtitle, thus leading to lower word fixation probabilities.

8.2.6 Median pupil size

The final dependent variable that will be discussed here is median pupil size. It is the first time a pupillometric measurement has been used to measure of processing effort experienced while viewing subtitled media. As hypotheses [9a] and [9b] state, it was expected that median pupil size would be larger when there was pop-up gloss on screen and when there was a two-line subtitle on screen.

As the data in Figure 8.13 illustrates, there was no difference in the average median pupil size when one-line subtitles were on screen whether or not there was a pop-up gloss on screen, contrary to our expectations. This stands in contrast with two-line subtitles, where the pupil size was larger when there was a pop-up gloss present. While there was a significant main effect for subtitle length, with a larger median pupil size evident in two-line subtitles than one-line subtitles ($M= 4.32\text{mm}$ vs. 4.25mm), $F(1,12) = 24.2$ $p<0.001$ $SS=0.07$ $MSe=0.01$, there was no significant effect shown for the subtitle type ($F(1,12) = 1.41$ $p>0.05$ $SS=0.01$ $MSe=0.01$), although the median pupil size was slightly larger in instances where there was a pop-up gloss present ($M=4.3\text{mm}$ vs. 4.27mm).

There was a significant interaction between both factors ($F(1,12) = 5.32$ $p < 0.05$ $SS = 0.01$ $MSe = 0.01$), with the effect of pop-up gloss on mean pupil size largest when two-line subtitles were watched. This suggests that more processing effort was required for two-line subtitles than one-line subtitles and that the presence of pop-up gloss evokes more processing effort when processing a two-line subtitle; while the processing effort required when a one-line subtitle was on screen was not significantly different in the gloss and solo conditions.

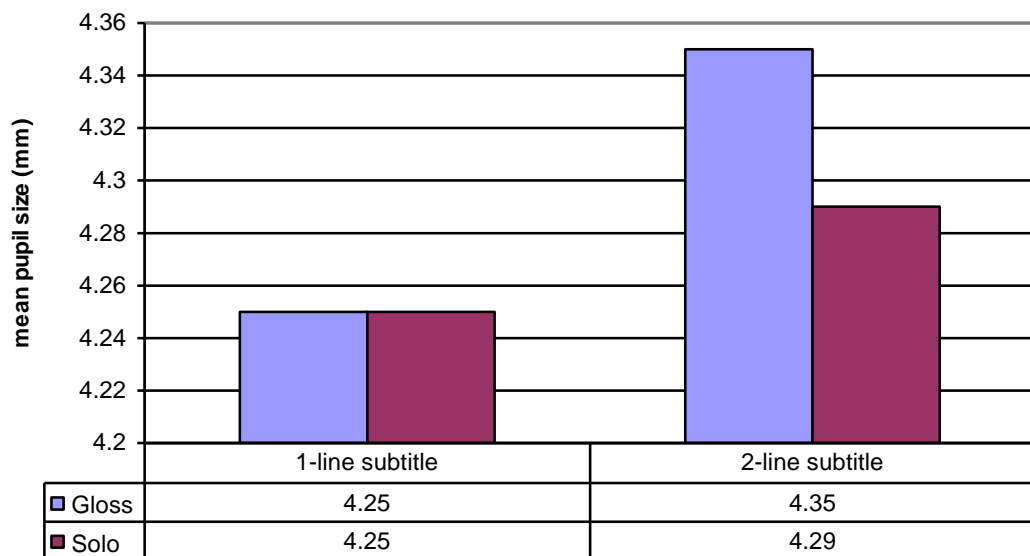


Figure 8.13. Mean median pupil size according to experiment condition

8.2.7 Summary of fixation-based and pupillometric results

- The data shows that the percentage of subtitles skipped was significantly higher when the subtitle on screen was a one-line subtitle, or when pop-up gloss was used. The percentage of subtitles skipped was at its highest when the subtitle was one line and was accompanied by a pop-up gloss.
- The percentage gaze time spent in subtitles was significantly lower when it was a one-line subtitle, and also when there was a pop-up gloss present at the same time as a subtitle.
- The mean fixation duration in subtitles was longer in one-line subtitles than two-line subtitles, and when there was no pop-up gloss on screen. The mean fixation duration was the longest when there were one-line subtitles on screen and no pop-up gloss present.

- The probability of a subject fixating on each word in the subtitle at least once was almost equal in one and two-line subtitles when there was no pop-up gloss present. However, if instances where there was a pop-up gloss present are included, the difference becomes significant. The presence of pop-up gloss coincided with a reduced word fixation probability. The likelihood of fixating on each word in the subtitle at least once was at its lowest in two-line subtitles on screen at the same time as a pop-up gloss.
- The mean pupil size did not differ when there was a one-line subtitle on screen whether or not there was a pop-up gloss present. However, it was larger in two-line subtitles than one-line subtitles, and was at its largest when a two-line subtitle was on-screen with a pop-up gloss.

8.3 Summary

Concerning PCEs, the presence of pop-up gloss coincided with an increased number of correct answers to CVNC questions in the present study, suggesting increased PCEs related to CVNCs when a pop-up gloss was used. The number of correct answers from subjects who saw the excerpts with pop-up gloss was significantly higher than from those who saw it without pop-up gloss. It was also found that subjects who watched the excerpts including pop-up gloss had a higher level of confidence in the accuracy of the information they retained regarding CVNCs, reflected by the higher number of subjects who saw the excerpts with pop-up gloss declaring that they understood the CVNC. The presence of pop-up gloss did not result in a significant difference in the number of subjects who experience PCEs relating to subtitles. However, two of the three instances where there was a relatively large difference (i.e. four subjects or more) between the number of subjects who got the subtitle question correct in the treatment and normal condition involved subtitles that were displayed while there was a pop-up gloss on screen.

In terms of processing effort, the results above indicate that the use of pop-up gloss did indeed have an impact on the amount of processing effort required to watch the excerpts. On a subjective level, more subjects watching the excerpt with pop-up gloss stated that they found the subtitle speed fast or too fast than subjects watching the same excerpt without pop-up gloss. This was despite the

subtitle speed being the same in both conditions. This can be explained by the presence of extra text when there was a pop-up gloss present requiring more processing effort be expended to read all of the on-screen text.

The fixation-based data also revealed some differences in the processing effort required depending on whether there was pop-up gloss present or not, and whether a one or two-line subtitle was on screen. When there was a pop-up gloss on screen, the number of subtitles skipped was significantly higher, while gaze time in the subtitle area was significantly lower, along with the word fixation probability and the mean fixation duration.

The percentage of one-line subtitles skipped was higher than two-line subtitles skipped, which was expected. The hypothesis that the percentage of skipped subtitles would also be higher when a pop-up gloss was present was also accepted. This was most evident in one-line subtitles, where the difference between subtitles skipped when a gloss was on-screen ($M=29.46\%$) and when there was no gloss ($M=5.51\%$) was larger at roughly 25%, than the difference between the same conditions in two-line subtitles at just over 4%. This data implies that more processing effort was required when a pop-up gloss was present than when there was none, as the amount of processing time dedicated to the subtitles was lower when there was a pop-up gloss present, indicated by the number of subtitles that were skipped.

The percentage of gaze time spent in subtitles was a further indication that the presence of pop-up gloss placed higher demands on subjects as far as processing effort was concerned. Whether or not a pop-up gloss was present, more time was spent in the subtitle area in two-line subtitles than one-line subtitles, indicating that more processing time was required for two-line subtitles, which supports the view that two-line subtitles are more semantically complex. When a pop-up gloss was on screen less time was spent in the subtitle area, irrespective of subtitle length. This may be related to the cue times for the subtitles remaining the same independent of the presence of pop-up gloss, resulting in the necessity for subjects to divide their attention more, plausibly increasing the processing effort required.

For both one and two-line subtitles the mean fixation duration was shorter when there was a pop-up gloss on screen. It was longest when there was a one-line subtitle on-screen with no gloss. In all other conditions, mean fixation durations were within 10 ms of each other. These results suggest that the presence of the pop-up gloss led to subjects spending less time in the subtitle area with each fixation, again possibly because of the need to allocate processing effort not only to the text in the subtitle and the image, but also to the text in the pop-up gloss.

While the mean word fixation probability was significantly lower when there was a pop-up gloss on screen, there was no significant difference between one and two-line subtitle conditions when there was no pop-up gloss on screen. This can be accounted for by the fact that the timing rules for subtitles tend to take subtitle length into account when deciding how long a subtitle remains on screen, with two-line subtitles remaining on screen for longer, allowing for more fixations the longer a subtitle is on screen. As the subtitle speed is not altered depending on the presence of pop-up gloss, when pop-up gloss appears on screen there is less time to fixate and process the subtitle, thus leading to lower word fixation probabilities.

The use of pupillometry in the current experiment was a novel aspect of the methodology, and as such, it was not known whether this form of measurement would be suited to measuring processing load with subtitled AV content. The results revealed that median pupil size was significantly larger for two-line subtitles than one-line subtitles. This was to be expected considering that the complexity of two-line subtitles tends to be higher than that of one-line subtitles (d'Ydewalle and de Bruycker 2007). There was no significant effect shown for the presence of pop-up gloss, which goes against intuition and the subjective measure of processing effort found in the questionnaire, and the fixation-based measures gathered with the eye tracker. However, there was a significant interaction between subtitle length and the presence of pop-up gloss, which revealed that the presence of pop-up gloss with a two-line subtitle had a significant effect on the median pupil size. Median pupil size was the largest in the gloss/two-line condition, which suggests that it was the gloss/two-line condition that required the most processing effort from subjects. On the other

hand, for one-line subtitles, there was no difference between the median pupil size in the gloss or solo conditions, indicating that the presence of pop-up gloss does not significantly increase the amount of processing effort experienced when a one-line subtitle is on screen.

In conclusion, the presence of pop-up gloss affected subjects in terms of both PCEs and processing effort. The lower amount of processing effort allocated to the subtitles, indicated by the fixation-based data; and the increased amount of processing effort required when pop-up gloss was present, indicated by the subjective questionnaire and pupillometric data, may be the cause for reduced PCEs relating to the subtitles experienced by subjects in some situations where a pop-up gloss was on-screen. The heat map images in Figure 8.14 display one instance where no fixations on either the subtitle or the CVNC were recorded from subjects in the treatment condition.

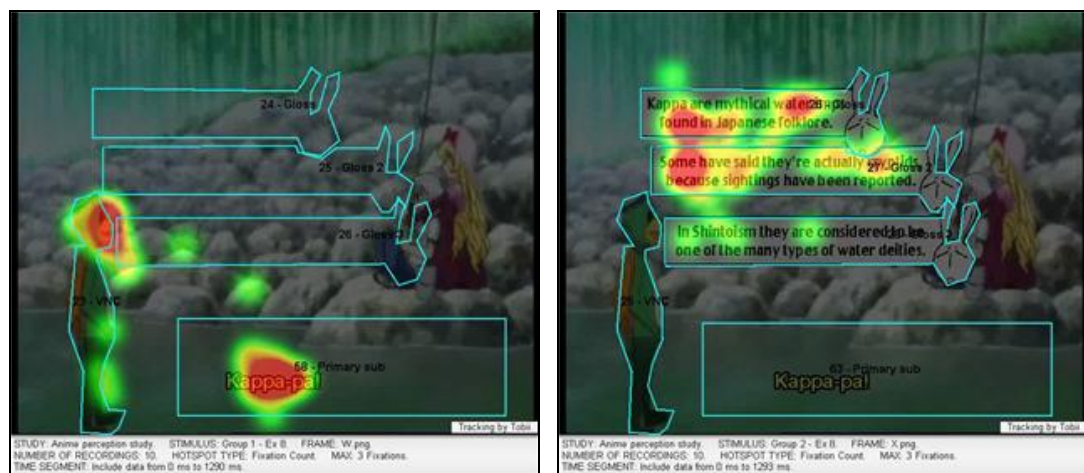


Figure 8.14. Fixation heatmap for same scene in normal (left) and treatment (right) conditions⁴⁰ (*Paniponi Dash!* 2007)

The red and yellow areas, which indicate parts of the image where the number of fixations was high, are all centred around the pop-up gloss in the treatment condition, while in the normal condition where no pop-up is used, the fixations are focused on the CVNC and the subtitle. Despite the lack of fixations in the subtitle area or the CVNC area, more subjects in the condition with the pop-up gloss present answered the CVNC question correctly in this instance (see Table 7.5, Excerpt 8 for content of pop-up gloss).

⁴⁰ Refer to Figure 7.1 to see the images as they were seen by the subjects

CHAPTER 9. **C**ONCLUSION

The empirical analysis reported in the previous chapter provided a rich description of the perception of the excerpts used in the experiment. This concluding chapter brings together the theoretical arguments and empirical analyses of the previous chapters by revisiting the research questions. The main contributions of the thesis, its limitations, and possible avenues for future research in the field of AVT perception studies are discussed.

9.1 Research questions revisited

Since the introduction of DVD technology in the 1990s, it has been possible for commercial producers of DVD to include multiple dubbing tracks and subtitle streams on one disc. This, coupled with the innovative use of abusive subtitle procedures by amateur and fan subtitlers on the internet, has led to the current situation where viewers of AV content are able to access an ever-broader variety of subtitles. The fact that these abusive procedures contravene current subtitling norms has not hampered their adoption by some commercial DVD distributors. Nonetheless, there is a striking paucity of empirical research into the effects that these procedures have on viewer perception of AV content.

Against this background, this thesis set out to empirically investigate the effects that one such abusive procedure, the pop-up gloss, has on viewer perception of excerpts from a subtitled TV anime on DVD. Two research questions thus constituted the focus of this research:

1. What effect does the presence of pop-up gloss have on the perception of subtitled TV anime?
2. Can pupillometric data be used to accurately measure the processing effort experienced while viewing subtitled AV content?

In order to answer the above questions, the thesis took as its starting point the notion of abusive subtitles (Chapter 2). The motivation behind their use on a commercially released TV anime DVD was discussed, and pop-up gloss was selected as a demonstrable abusive procedure for investigation in the current study.

The source text on which the thesis focuses, TV anime on DVD, was then described (Chapter 3). Some of the distinctive aspects of TV anime as a source text were discussed, and the two aspects of the AV content, subtitles and CVNCs, that would be the focus of the perception study were explained. A definition for the term CVNC was provided, where it is described as a nonverbal item appearing in the image of an audiovisual text whose function and connotation in the source text involve a translation problem in its transfer to the target text whenever this problem is a product of the non-existence of the item or of its different intertextual status in the cultural system of the target viewers.

Throughout the thesis, the perception of AV content was considered through the lens of relevance theory (Chapter 4). Its two criteria for judging the relevance of an input for an individual, PCEs and processing effort (Wilson and Sperber 2002), provided the conceptual tools with which to analyse perception. Two data collection methods were used to gather data on PCEs and processing effort, questionnaire and eye tracker. The eye tracker was chosen as a suitable tool for gathering data on processing effort (Chapter 5). Two types of data it provides, fixation-based and pupillometric data, were explained in detail, and the specific measurements that would be used for each type of data revealed.

A critical review was made of prior relevant perception studies in the field of AVT (Chapter 6). This literature review, along with a pilot study, informed the methodological considerations and experiment design concerning this study (Chapter 7). These considerations were met by designing an experiment where 20 subjects viewed a mixture of subtitled excerpts, half with the abusive procedure (i.e. pop-up gloss) and half without, obtained from a commercially produced TV anime DVD. While viewing the excerpts, subjects' eye movements and pupil size were simultaneously recorded with an eye tracker. After watching the excerpts, the subjects answered a questionnaire consisting of four questions on each excerpt.

The analysis of the data generated from the experiment provided a representation of the PCEs and processing effort experienced by subjects, and the effects that the presence of pop-up gloss had on the same (Chapter 8). By

combining this data, the findings in relation to the research questions can now be put forward.

9.1.1 Question 1

What effect does the presence of pop-up gloss have on the perception of subtitled TV anime?

The data presented in the previous chapter (Chapter 8) provides empirical evidence to suggest that the presence of pop-up gloss does indeed influence viewer perception of subtitled TV anime. More specifically, the number of subjects who experienced greater perceived and actual PCEs relating to CVNCs was significantly higher in the treatment condition (i.e. subjects who watched the excerpt with pop-up gloss) than in the normal condition (i.e. subjects who watched the same excerpt without pop-up gloss). This indicates that the use of this abusive subtitling procedure does indeed improve viewer understanding of culturally marked items in the AV content. An unexpected result further supporting the use of the abusive procedure is the lack of any significant difference between experiment conditions concerning PCEs relating to subtitles. However, in some instances where pop-up glosses were on screen at the same time as a subtitle, there were no fixations recorded in the subtitle area, indicating that some subjects spent no time reading the subtitle.

The subjective data from subjects' self assessment of processing effort obtained after watching the excerpts was combined with the more objective fixation-based and pupillometric measurements taken with the eye tracker. Data collected using both questionnaire and eye tracker suggest that the use of pop-up gloss does increase the processing effort experienced by viewers of subtitled TV anime. One indication of this is that a greater number of subjects perceived the subtitle speed as faster in the treatment condition than in the normal condition, despite the subtitle speed being the same.

The fixation-based data also suggests that more processing effort is required when a pop-up gloss is on screen together with a subtitle, resulting in less processing time being allocated to the subtitle. There was a higher number of subtitles skipped when a pop-up gloss was present, as well as a lower

percentage of gaze time in the subtitle area, mean fixation duration and word fixation probability.

The pupillometric results imply that two-line subtitles require more processing effort than one-line subtitles. When a one-line subtitle was on screen, the presence of pop-up gloss did not have any significant effect on the pupil size, indicating that there was no significant effect on the processing effort experienced by subjects according to this measurement. However, there was a significant interaction between factors, with the median pupil size being significantly larger when a two-line subtitle and pop-up gloss were both present.

Considering the data on the effects of pop-up gloss on the perception of subtitled excerpts in terms of PCEs and processing effort, several implications become evident. Taking the subjects' self assessment from the consecutively gathered questionnaire data alone, it would appear that a trade-off takes place between the higher number of perceived PCEs relating to CVNCs when the AV content is watched with pop-up gloss, with subjects finding the speed of subtitles faster, indicating that increased processing effort is required.

On the other hand, the more objective measurement of processing effort provided by pupillometric data suggests that it is only when a two-line subtitle is on screen that the presence of pop-up gloss places more demands on viewers' processing. Taking these results alone would suggest that the use of pop-up gloss when a one-line subtitle is on screen could improve the relevance of a translation, providing more PCEs for the same, or at least not a significantly higher, level of processing effort.

However, consideration of the fixation-based data casts doubt on any prospect that the use of pop-up gloss with one-line subtitles would improve the relevance of subtitled AV content. While processing effort may not be significantly different in relation to the presence of pop-up gloss on screen at the same time as a one-line subtitle, fixation-based data suggests that the apportioning of processing effort is. With almost 30% of one-line subtitles not being fixated upon at all when a pop-up gloss is present (compared with just over 5% when no pop-up gloss is present) there is bound to be a lower level of overall PCEs

related to one-line subtitles that are on screen at the same time as a pop-up gloss.

9.1.2 Question 2

Can pupillometric data be used to accurately measure the processing effort experienced while viewing subtitled AV content?

The pupillometric data suggests that two-line subtitles required significantly more processing effort than one-line subtitles. This is a plausible result, as it concurs with the fixation-based measurements, and previous research (d'Ydewalle and de Bruycker 2007), which indicates that two-line subtitles require more processing effort. This may be put down to the increased complexity of two-line subtitles and the semiotic redundancy of one-line subtitles mentioned previously.

Further consideration should be given to the implication that there is no difference in the processing effort required when a one-line subtitle is on screen in relation to the presence of pop-up gloss mentioned in the previous section. This lack of difference goes against intuition, particularly when the percentage of skipped one-line subtitles when a pop-up gloss is on screen is considered. This could place some doubt on the applicability of pupillometric data for measuring the processing effort experienced when viewing subtitled AV content. However, the median pupil size was significantly larger when a pop-up gloss was on screen with a two-line subtitle, indicating an increase in the effort required by subjects, and while the percentage of two-line subtitles skipped was higher when a pop-up gloss was present, it was still relatively low at 4.46%.

The difference in data for one and two-line subtitles could be explained as follows. The semiotic redundancy of one-line subtitles plausibly resulted in processing effort being allocated to the pop-up gloss rather than being “wasted” on the one-line subtitles, fully in line with the principles of minimum processing effort for maximum informational gain propounded by relevance theory. This would explain both the higher percentage of skipped subtitles (29.46%) and equal median pupil size when a one-line subtitle was on screen with a pop-up gloss.

On the other hand, the lack of semiotic redundancy in two-line subtitles meant that the viewer could not afford to ignore the subtitles as often, and therefore needed to apportion processing effort to both the subtitle and the pop-up gloss, resulting in an increase in processing effort required. This may have resulted in a lower percentage of skipped subtitles (4.46%) than in the case of one-line subtitles with pop-up gloss, and an increase in the median pupil size.

The conclusion that may be drawn from the pupillometric data on one-line and two-line subtitles when pop-up gloss is present is that the pupillometric measure of processing effort accurately reflected the effort experienced by the subjects in the current study. This is an indication that pupillometric data can indeed be used to accurately measure the processing effort experienced while viewing subtitled AV content. It forms the beginning of an interesting inquiry into how the brain works with subtitled AV content, which will require further study for verification.

9.2 Strengths and limitations of the present study

Against the lack of systematic data analysis supporting the theoretical arguments surrounding pop-up gloss and other abusive subtitling procedures, this thesis succeeds in putting forward a conceptual and methodological framework substantiated by empirical evidence for analysing the effects of pop-up gloss on the perception of subtitled TV anime. It is hoped that the methodology employed in this research will prove useful to determine the effects of other AV translation procedures and modalities on viewer perception.

A more particular aspect of the methodology that presented some difficulties was the use of PCEs in the experiment. In order to measure PCEs using questionnaire, “correct” answers had to be decided to form the basis of judgement for whether a cognitive effect could be considered positive or not. This was relatively easy in the case of the subtitles, where the answers were in the subtitles themselves, but in the case of CVNCs it was more difficult. In this study, the answers for the CVNC were taken from the pop-up gloss, but in some cases this decision may be flawed, as it automatically assumes that the opinion

of the writer of the pop-up gloss is correct⁴¹. While this method is prescriptive, it was unavoidably so, if a relevance theoretical framework was to be used in such a study.

The measurement of PCEs also presented some difficulties. While processing effort could be measured continuously while subjects watched the excerpts, it was not possible to do so with PCEs without asking subjects a question on each subtitle, and about each CVNC, which would likely have resulted in an increase in fatigue-related negative influences on the data.

The final point regarding PCEs and the design of the experiment, concerns future studies using a similar questionnaire design. It is suggested that the PCE question is asked about a subtitle that is on screen at the same time as the translation procedure that is being tested. This would perhaps provide more of an insight into the effect of the procedure than the design in the present study, where some subtitles that were asked about were on screen with pop-up gloss and some without.

The measurement of processing effort was difficult, as the subjects were not supposed to know the focus of the study. For that reason, direct questions about the processing effort they experienced were avoided. Instead, subjects were asked how they perceived the speed of the subtitle set in each excerpt. While this indicated the subtitle speed as assessed subjectively by subjects, it does not necessarily translate into a measurement of processing effort required. For the purpose of the current study, the assumption was made that if a subject finds a subtitle set is quick, then they will require more processing effort to view it. However, this assumption has not been tested, and could be refined for future studies.

The research also brought to light some of the consequences in using the eye tracker as a data collection tool. The application of a new tool, in the form of the *Tobii 1750* eye tracker, opens up a new approach to investigating the perception of subtitled AV content. The study also contributes towards gaining

⁴¹ However, in the current study, the accuracy of the writers was assumed only after the content of each pop-up gloss was checked for accuracy.

a clearer understanding of the link between vision and the brain, through the investigation of abusive subtitle procedures applied to AV content, resulting in possible applications in the field of AVT, film studies and cognitive psychology. Although the *Tobii 1750* eye tracker has a less steep learning curve in comparison to other head-mounted eye trackers, there are still many practical issues involved in its use. Prime among these is the vast quantity of data that the tool generates, and therefore the analysis of that data. The lack of software to deal automatically with the pupillometric data makes it particularly time consuming, a problem that is not remedied with the updated software package *Tobii Studio*.

As mentioned in Section 7.5.3.1, the inability to pause the excerpts when viewing them on the eye tracker may have corrupted the realism of the experimental setting. Although this was unplanned, and unavoidable if the eye tracker was to be used, it should still be of interest to the producers of AV content to gather data on what happens when pop-up gloss is shown at the same time as other information on the screen and watched without interruption. This data could be used for the development of future strategies relating to pop-up gloss, and indeed other abusive or experimental procedures that will be made possible by media developments in the future.

Concerning pupillometric data, while the results obtained using the pupil size data do indicate that pupillometry could be used as a means of measuring processing effort experienced when watching subtitled AV content, the lack of control over a number of factors were likely to have had an impact on the pupillometric data. While this was an unfortunate necessity in the present study where commercially released AV content was used, for future research a more controlled experiment, where all of the factors of the AV content are controlled by the researcher may be advantageous and provide more distinctive findings. For example, to limit the effects that content from various semiotic channels has on pupil size, and individual differences in pupil size, a longitudinal study could be carried out, where the exact same excerpt is shown to the same subjects with two different subtitle sets after a period of several months. By doing so, the likelihood of subjects remembering the content may be reduced, and the information coming from each semiotic channel will be the same, except for the

subtitles being used. The results of such an experiment with these controls could then be used to formulate a list of recommendations for commercial distributors about which circumstances are favourable for the use of pop-up gloss, and other subtitle procedures that elicited a higher level of processing effort.

Considerable thought was given to the decision about whether statistical testing would be carried out on the data, and if so, which tests would be used. Finally an expert⁴² was consulted to check the statistical testing employed in order to assure the validity of the statistical tests used. Here the small sample size did have an effect on the explanatory power of the tests, meaning that the results could not be readily generalised for an entire population. However, the use of statistical tests still provided an objective measure against which the significance of a result could be measured. The increased use of data collection tools such as the eye tracker where large amounts of data are gathered and need to be analysed provides a requirement for some training on statistical methods on the part of researchers in the context of empirical research in TS.

9.3 Directions for future work

The methodology developed in this study is highly adaptive and could be applied to research into other areas of subtitling and AVT in general. An example of one area where it could be particularly useful is the domain of audio description (AD), which is currently attracting research interests in AVT. AD is the process where an additional narrative level of sound is added to AV content describing the actions that take place in the visual channel, and is a means of making AV content accessible to the blind and visually impaired (Benecke 2004, Orero 2008). There is currently a debate among some researchers as the degree to which AD commentary should be subjective (Orero and Vercauteren 2008). The methodology from the present study could be adapted to such a study as follows. Consideration should be given to the way in which positive cognitive effects and processing effort will be measured. If the aspect of interest is emotions, the PCEs relating to these can be determined by asking

⁴² The researcher would like to thank Elisa Alonso Sanz, Consultant from Ernst and Young Luxembourg for her assistance with the statistics.

subjects to describe what emotion they think the character is showing. Processing effort can again be measured subjectively here, asking subjects which version they find to be more strenuous, and while eye tracking may not be used in this case (if blind or visually impaired subjects are to be used), there are other methods of measuring processing effort which may be used to complement subjective questionnaire or interview data. One possibility is the use of *event-related brain potentials* (ERPs), which measure the electrical fields around the head and were found to be a reliable means of measuring processing effort while reading texts on computer by Schultheis and Jameson (2004).

Returning to the eye tracker, while the study only used the diagnostic feature of the eye tracker, if a reliable simultaneous measure of cognitive load could be determined by using pupillometry, it would open several avenues for interactive features. As the price of eye tracking technology decreases, it is plausible that it will become more widely available at lower prices. Also, as the accuracy and range of integrated eye trackers improve, the notion of adaptive subtitles may become a reality. Adaptive subtitles are subtitles that would adapt to viewers' individual needs using a combination of preset viewer preferences and simultaneously gathered physiological data from the viewer. With television's transition to digital, and IPTV (internet protocol television)⁴³ on the horizon, the provision of these adaptive subtitles should be made easier. The online storage available to an IPTV provider would allow them to store many more subtitle tracks for the content they offer than is afforded by DVD, or Blu-ray. As such, broadcasters could provide viewers with a huge number of subtitle streams, for example, a highly domesticated subtitle stream for average viewers, more foreignised streams for fans, or even streams including pop-up glosses to explain a variety of items for viewers with a particular cultural interest. If combined with an accurate fine-grained simultaneous algorithm for analysing pupillometric data, the processing effort experienced by the viewer could be determined, and according to a preset boundary, the subtitle speed could be decreased or content excluded when the effort required is "too high", and conversely, the subtitle speed increased, or pop-up glosses included when the effort required is "low".

⁴³ IPTV, or internet protocol television is television provided over the internet, using the same technology.

These are just two examples of the many applications of the research presented in this thesis to future research. It is hoped that this thesis succeeded in providing a basis from which fixation-based and pupillometric data gathered using eye tracker can be combined with questionnaire data to provide a holistic means of measuring the effects of abusive translation strategies on viewer perception of subtitled AV content. There are many more abusive procedures to be investigated, and many more AVT modalities where eye tracking can offer a new outlook on viewer perception. The recent surge in interest in investigating the perception of translated AV content using eye tracker evinced by events such as the “Seminar on Eye-Tracking and Audiovisual Translation” mentioned in the introduction, and the forthcoming volume on eye tracking and AVT edited by Elisa Perego and Géry d’Ydewalle are promising indicators that this research area will be here for some time to come.

References

ADV Films. 2005. United States Patent and Trademark Office. *Vid-notes*. 78271566.

Ahern, S. and Beatty, J. 1979. Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science*. 205, pp1289-1292.

Aixelá, J.F. 1996. Culture-specific items in translation. *IN: Alvarez, R. and Carmen-Africa Vidal, M. (eds.) Translation, Power, Subversion*. Clevedon: Multilingual Matters, pp 52-78.

Alves, F. 2003. *Triangulating Translation: Perspectives in Process Oriented Research*. Philadelphia: John Benjamins.

Antonini, R. 2005. The perception of subtitled humour in Italy: An empirical study. *Humor, Journal of Humor Research*. 18 (2), pp 209-225.

Antonini, R. 2007. SAT, BAT, spirit biscuits and the third amendment: What Italians make of cultural references in dubbed texts. *IN: Gambier, Y., Schlesinger, M., and Stolze, R. (eds.) Doubts and Directions in Translation Studies: Selected Contributions from the EST Congress, Lisbon 2004*. Amsterdam: John Benjamins, pp 153-168.

Antonini, R. and Chiaro, D. 2005. The perception of subtitled humour in Italy: An empirical study. *IN: Bondi, M. and Maxwell, N. (eds.) Cross-Cultural Encounters: Linguistic Perspectives*. Rome: Officina Edizioni, pp 33-44.

Antonini, R. and Chiaro, D. 2009. The perception of dubbing by Italian audiences. *IN: Anderman, G. and Díaz Cintas, J. (eds.) Audiovisual Translation: Language Transfer on Screen*. Basingstoke: Palgrave MacMillan, pp 97-114.

Appiah, K.A. 2000. Thick translation. *IN: Venuti, L. (ed.) The Translation Studies Reader*. New York: Routledge, pp 417-429.

Austin, T. 2005. Seeing, feeling, knowing: A case study of audience perspectives on screen documentary. *Particip@tions* [Online]. 2 (1). Available from:

<http://www.participations.org/volume%202/issue%201/2_01_austin.htm>
[Accessed 22 August 08]

Baker R. G. 1982. *Monitoring Eye-movements While Watching Subtitled Television Programmes – a Feasibility Study*. London: Independent Broadcasting Authority.

Baumgarten, N. 2008. Yeah, that's it! Verbal reference to visual information in film texts and film translations. *Meta*. 53 (1), pp 6 – 25.

Beatty, J. 1982. Task-evoked pupillary responses, processing load and the structure of processing resources. *Psychological Bulletin*. 91 (2), pp 276-292.

- Beatty, J. and Lucero-Wagner, B. 2000. Pupillary system. *IN: Cacioppo, J.T., Tassinary, L.G., and Berntson, G. (eds.) Handbook of Psychophysiology. 2nd ed.* Cambridge: Cambridge University Press, pp 142-163.
- Beckman, R. 2008. An out-of-character role for subtitles. *Washington Post* [Online], 16 November 2008. Available from:
<<http://www.washingtonpost.com/wp-dyn/content/article/2008/11/14/AR2008111400700.html>>
[Accessed 26 January 09]
- Benecke, B. 2004. Audio-description. *Meta*. 49 (1), pp 78-80.
- Beymer, D., Orton, Z., and Russell, D.M. 2007. An eye tracking study of how pictures influence online reading. *Interact*. 2, pp 456-460.
- Bogucki, L. 2004. The constraint of relevance in subtitling. *The Journal of Specialised Translation* [Online], 1. Available from:
<http://www.jostrans.org/issue01/art_bogucki_en.php>
[Accessed 22 August 08]
- Brown, S. (ed.) 2006. *Cinema Anime: Critical Engagements with Japanese Animation*. New York: Palgrave Macmillan.
- Bucaria, C. 2005. *The Perception of the Non-Verbal: How Italians React to What They see but isn't Explained*. EU High Level Conference Series – Challenges of Multidimensional Translation, Saarbrücken, Germany, 2-6 May 2005.
- Bucaria, C. and Chiaro, D. 2007. End user perception of screen translation: the case of Italian dubbing. *Tradterm*. 13(1), pp 91–118.
- Caffrey, C. 2007. *Innovative Insanity? “Abusive” Subtitling Strategies and their Effects on Viewers’ Reading Times*. Intermedia Conference, Łódź, Poland, 13-14 April 2007.
- Caffrey, C. 2008a. Viewer perception of visual nonverbal cues in subtitled TV Anime. *European Journal of English Studies*. 12(2), pp163 – 178.
- Caffrey, C. 2008b. *Relevant Abuse? Effects of Abusive Subtitling Strategies on Viewer Perception of TV Anime*. Audiovisual Translation: Multidisciplinary Approaches Conference, Montpellier, France, 19 -20 June 2008.
- Caffrey, C. 2008c. Using pupillometric, fixation-based and subjective measures to measure the processing effort experienced when viewing subtitled TV anime with pop-up gloss. *IN: Göpferich, S., Jakobsen, A.L. and Mees, I.M. (eds.) Looking at Eyes. Eye-Tracking Studies of Reading and Translation Processing. Copenhagen Studies in Language*, 36, pp125 – 144.
- Carrol, M. 2004. Subtitling: Changing standards for new media? *The Globalisation Insider* [Online]. 9. Available from:
<http://www.lisa.org/globalizationinsider/2004/09/subtitling_chan.html>

[Accessed 29 September 08]

Castro Roig, X. 2001. Solo ante el subtítulo. Experiencias de un subtitulador. *IN: Lorenzo García, L., and Pereira Rodríguez, A.M. (eds.) Traducción Subordinada (II). El Subtitulado*. Vigo: Servicio de Publicacións de la Universidad de Vigo, pp 19-24.

Chaume Varela, F. 1997. Translating non-verbal information in dubbing. *IN: Poyatos, F. (ed.) Nonverbal Communication and Translation: New Perspectives and Challenges in Literature, Interpretation, and the Media*. Amsterdam: John Benjamins, pp 315-326.

Chesterman, A. 2000. Empirical research methods in Translation Studies. *Erikoiskielet Ja Käännösteoria (VAKKI-symposiumi XX)*. 27, pp 9-22.

Chesterman, A. 2005. Towards consilience? *IN: Aijmer, K. and Alvstad, C. (eds.) New Tendencies in Translation Studies*. Göteborg: Acta Universitatis Gothoburgensis, pp 19-27.

Chiaro, D. 2004. Investigating the perception of translated verbally expressed humour on Italian TV. *ESP Across Cultures*. 1, pp 35-52.

Chiaro, D. 2007. The effect of translation on humour response: The case of dubbed comedy in Italy. *IN: Gambier, Y., Schlesinger, M., and Stolze, R. (eds.) Doubts and Directions in Translation Studies: Selected Contributions from the EST Congress, Lisbon 2004*. Amsterdam: John Benjamins, pp 137-152.

Conyngham, G. 2007. *Statistical Tests and Statistical Significance*. Lecture notes distributed in postgraduate research module.

Cubbison, L. 2005. Anime fans, DVDs, and the authentic text. *The Velvet Light Trap*. 56, pp 45-57.

Dean, P. 2007. DVD add-ons or by-gones? *Convergence: The International Journal of Research into New Media Technologies*. 13 (2), pp 119-128.

Delabastita, D. 1996. Introduction. *The Translator*. 2 (2), pp 127-139.

de Linde, Z. and Kay, N. 1999a. *The Semiotics of Subtitling*. Manchester: St. Jerome.

de Linde, Z. and Kay, N. 1999b. Processing subtitles and film images: Hearing vs. deaf viewers. *The Translator*. 5(1), pp 45-60.

Díaz Cintas, J. 2004a. In search of a theoretical framework for the study of audiovisual translation. *IN: Orero, P. (ed.) Topics in Audiovisual Translation*. Amsterdam: John Benjamins, pp 21-34.

Díaz Cintas, J. 2004b. Subtitling: the long road to academic acknowledgment. *The Journal of Specialised Translation* [Online]. 1. Available from: <http://www.jostrans.org/issue01/art_diaz_cintas.php> [Accessed 22 August 08]

- Díaz Cintas, J. 2005. El subtitulado y los avances tecnológicos. *IN: Merino, R., Santamaria, J.M. and Parajes, E. (eds.) Traspases Culturales: Literatura, Cine, Traducción 4*. Vitoria: Universidad del País Vasco, pp 155-175.
- Díaz Cintas, J. and Muñoz Sánchez, P. 2006. Fansubs: Audiovisual translation in an amateur environment. *The Journal of Specialised Translation* [Online]. 6. Available from:
<http://www.jostrans.org/issue06/art_diaz_munoz.php>
[Accessed 22 August 08]
- Díaz Cintas, J. and Remael, A. 2007. *Audiovisual Translation: Subtitling*. Manchester: St. Jerome Publishing.
- Duchowski, A. T. 2003. *Eye Tracking Methodology: Theory and Practice*. New York, Springer.
- d'Ydewalle, G. and de Bruycker, W. 2007. Eye movements of children and adults while reading television subtitles. *European Psychologist*.12 (3), pp 196-205.
- d'Ydewalle, G. and Gielen, I. 1992. Attention allocation with overlapping sound, image, and text. *IN: Rayner, K. (ed.) Eye Movements and Visual Cognition: Scene Perception and Reading*. New York: Springer-Verlag, pp 415-427.
- d'Ydewalle, G., Praet, C., Verfaillie, K., and Van Rensbergen, J. 1991. Watching subtitled television: Automatic reading behaviour. *Communication Research*. 18, pp 650-666.
- d'Ydewalle, G., Van Rensbergen, J., and Pollet, J. 1985. *Reading a Message When the Same Message is Auditorily Available in Another Language: The Case of Subtitling (Psych. Rep. No. 54)*. Leuven: Katholieke Universiteit Leuven.
- Fadem, B. 2008. *High-Yield Behavioral Science*. Maryland: Lippincott Williams and Wilkins.
- Ferrer Simó, M. R. 2005. Fansubs y scanlations: la influencia del aficionado en los criterios profesionales. *Puentes*. 6, pp 27-43.
- Fisher, D. F., Karsh, R., Breitenbach, F., and Barnette, B. D. 1983. Eye movements and picture recognition: Contribution or embellishment? *IN: Groner, R., Menz, C., Fisher, D. E., and Monty, R. A. (eds.) Eye Movements and Psychological Functions: International Views*. New Jersey: Erlbaum, pp 193-210.
- Frey, L.R., Botan, C.H., Friedman, P.G., Kreps, G.L. 1991. *Investigating Communication: An Introduction to Research Methods*. New Jersey: Prentice Hall.

- Fuentes Luque, A. 2003. An empirical approach to the reception of AV translated humour. A case study of the Marx brothers' Duck Soup. *The Translator*. 9(2), pp 293-306.
- Gambier, Y. 2003. Screen transadaptation: Perception and reception. *The Translator*. 9(2), pp 171-190.
- Gambier, Y. 2005. *Is Audiovisual Translation the Future of Translation Studies? Between Text and Image*. Updating Research in Screen Translation conference, Forlí, Italy, 27-29 October 2005.
- Gibbs, R.W. and Tendahl, M. 2006. Cognitive effort and effects in metaphor comprehension: Relevance theory and psycholinguistics. *Mind and Language*. 21, pp 379-403.
- Gielen, I. 1988a. *Het Verwerken van Ondertitels in Functie van de Klankband en Typen van Informatie in het Beeld*. Licentiaatverhandeling, Leuven University.
- Gielen, M. 1988b. *Perceptie en Ondertitels: De Parafoveale en Perifere Informatie Verwerking van Ondertitels*. Licentiaatverhandeling. Leuven University.
- Goldberg, J.H., Stimson, M.J., Lewenstein, M., Scott, N., and Wichansky, A.M. 2002. Eye tracking in web search tasks: design implications. *Proceedings of the Eye Tracking Research and Applications Symposium 2002*. New York: ACM, pp 51-58.
- Gottlieb, H. 1992. Subtitling – a new university discipline. *IN: Dollerup, C. and Loddegaard, A. (eds.) Teaching Translation and Interpreting*. Amsterdam: John Benjamins, pp 161-170.
- Gottlieb, H. 1994. Subtitling - people translating people. *IN: Dollerup, C. and Lindegaard, A. (eds.) Teaching Translation and Interpreting II: Insights, Aims, Visions*. Amsterdam: John Benjamins, pp 261-274.
- Gottlieb, H. 1995. Establishing a framework for a typology of subtitle reading strategies - DK Viewer reactions to deviations from subtitling standards. *Communication Audiovisuelle et Transferts Linguistiques. Translatio – Nouvelles de la FIT*. 14(3-4), pp 388-409.
- Gottlieb, H. 2005a. *Multidimensional Translation: Semantics turned Semiotics*. EU High Level Conference Series – Challenges of Multidimensional Translation, Saarbrücken, Germany, 2-6 May 2005.
- Gottlieb, H. 2005b. Texts, translation and subtitling – In theory, and in Denmark. *IN: Gottlieb, H. (ed.) Screen Translation. Eight Studies in Subtitling, Dubbing and Voice-over*. Copenhagen: University of Copenhagen, pp 1-40.
- Granka, L. and Rodden, K. 2006. *Incorporating Eyetracking into User Studies at Google* [Online]. Workshop position paper presented at CHI 2006, ACM. Available from:

- <www.amber-light.co.uk/CHI2006/GrankaRodden-EyeTrackingWorkshop.pdf>
[Accessed 20 December 08]
- Griffin, Z.M. 2001. *Eye Tracking for People Who Don't Care About Vision or How to Get More Dependent Features*. Lecture notes distributed in graduate research methods at Stanford University, 19th February 2001. Available from: <<http://www-psych.stanford.edu/~lera/290/EyeTrack.pdf>>
[Accessed 5 November 07]
- Gutt, E. A. 1991. *Translation and Relevance. Cognition and Context*. London: Basil Blackwell.
- He, E. and Kowler, E. 1992. The role of saccades in the perception of texture patterns. *Vision Research*. 32, pp 2151-2163.
- Henry, C. 2006. *Exploring the Hidden World of Japanese Anime Fansubbing. What Can Professionals Learn?* CTIS Monday Seminar Series, Manchester, England, 9 October 2006.
- Hermans, T. 2003. Cross-cultural translation studies as thick translation. *Bulletin of the School of African and Oriental Studies*. 66(3), pp 380 – 389.
- Hess, E. H. 1972. Pupillometrics: A method of studying mental, emotional, and sensory processes. *IN: Greenfield, N. S. and Sternbach, R. A. (eds.) Handbook of Psychophysiology*. New York: Holt, Rinehart and Winston, pp 491-531.
- Holub, R. 1984. *Reception Theory. A Critical Introduction*. London: Methuen.
- House, J. 1981. *A Model for Translation Quality Assessment*. Tübingen: Gunter Narr.
- House, J. 1997. *Translation Quality Assessment: A Model Revisited*. Tübingen: Gunter Narr.
- Howell, P. 2005. *Textual Procedures and Strategies in the Translation of Manga and Anime Dialogue*. Ph.D. thesis. University of Edinburgh.
- Howell, P. 2006. Character voice in anime subtitles. *Perspectives: Studies in Translatology*. 14 (4), pp 292-305.
- Hyönä, J., Tommola, J., and Alaja, A.M. 1995. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology*. 48 (a), pp598 – 612.
- Igarashi, R. 2008. *Ru's Tenth Annual Anime Fan Questionnaire 2007* [Online]. Available from: <<http://members.shaw.ca/ru.igarashi/Surveys/fans/fan0709/results.html>>
[Accessed 10 October 08]
- Iqbal, S. T., Adamczyk, D., Sam Zheng, X., and Bailey, B. P. 2005. Towards an index of opportunity: Understanding changes in mental workload during task

- execution. *Proceedings of CHI 2005, 2nd-7th April 2005, Oregon USA*, pp311-320.
- Ishinomori, S. 1988. *Japan Inc. An Introduction to Japanese Economics*. Translated from the Japanese 日本経済入門(1986) by Betsy Scheiner. London: University of California Press.
- Janisse, M.P. 1977. *Pupillometry: The Psychology of the Pupillary Response*. Washington: Hemisphere Publishing Corporation.
- JETRO. 2005. Japan animation industry trends. *JETRO Japan Economic Monthly* [Online]. Available from:
<www.jetro.go.jp/en/reports/market/pdf/2005_35_r.pdf>
[Accessed 5 May 08]
- Jaskanen, S. 1999. *On the Inside Track to Loserville, USA: Strategies Used in Translating Humour in Two Finnish Versions of Reality Bites*. Thesis. Helsinki University.
- Jensema, C.J. 1998. Viewer reaction to different television captioning speeds. *American Annals of the Deaf*. 143 (4), pp 318-324.
- Jensema, C. J., El Sharkawy, S., Sarma Danturthi, R., Burch, R. and Hsu, D. 2000a. Eye movement patterns of captioned television viewers. *American Annals of the Deaf*. 145 (3), pp 275-285.
- Jensema, C.J., Sarma Danturthi, R., and Burch, R. 2000b. Time spent viewing captions on television programs. *American Annals of the Deaf*. 145 (5), pp 464-468.
- Jungheim, N.O. 2000. Nonverbal behavior and refusals in Japanese anime: Sazae san. *Pragmatic Matters*. 2 (1), pp 9–10.
- Jungheim, N.O. 2004. Perception of gestures in Japanese media: The interpretation of refusals in two contexts. *IN: Jungheim, N. (ed.) メディアの異文化間の影響 [media no ibunkakan no eikyō, cross-cultural effects of media]*. Tokyo: The Research Institute of Aoyama Gakuin University.
- Jungheim, N.O. 2006. Learner and native speaker perspectives on a culturally-specific Japanese refusal gesture. *IRAL*. 44, pp 63 – 96.
- Jüngst, H. 2004. Japanese comics in Germany. *Perspectives: Studies in Translatology*. 12 (1), pp 83-105.
- Just, M.A. and Carpenter, A. 1993. The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*. 47 (2), pp 310-339.
- Kahneman, D., and Beatty, J. 1966. Pupil diameter and load on memory. *Science*. 154, pp 1583-1585.

- Karamitroglou, F. 1998. A proposed set of subtitling standards. *Translation journal* [Online]. 2 (2). Available from: <<http://accurapid.com/journal/04stndrd.htm>> [Accessed 20 November 08]
- Klingner, J., Kumar, R., and Hanrahan, P. 2008. Measuring the task-evoked pupillary response with a remote eye tracker. *Proceedings from ETRA 2008*, pp 69-72.
- Koolstra, C.M., Van der Voort, T.H.A., and d'Ydewalle, G. 1999. Lengthening the presentation time of subtitles on television: Effects on children's reading time and recognition. *Communications*. 24, pp 407-422.
- Kovačič, I. 1994. Relevance as a factor in subtitling reductions. *IN: Dollerup, C. and Lindegaard, A. (eds.) Teaching Translation and Interpreting 2*. Amsterdam: John Benjamins, pp 245-251.
- Kovačič, I. 1995. Reception of subtitles. The non-existent ideal viewer. *FIT Newsletter- Nouvelles de la FIT*. 14 (3-4), pp 376-383.
- Kramer, A.F. 1991. Physiological metrics of mental workload: A review of recent progress. *IN: Damos, D.L. (ed.) Multiple-Task Performance*. London: Taylor and Francis, pp 329-360.
- Krause, W. 1982. Eye fixations and three-term series problems, or: Is there evidence for task independent information units? *IN: Groner, R. and Fraisse, P. (eds.) Cognition and Eye Movements*. Berlin: Deutscher Verlag der Wissenschaften, pp 122-138.
- Kress, G. and Van Leeuwen, T.J. 1996. *Reading Images - The Grammar of Visual Design*. London: Routledge.
- Krüger, F. 2000. *Coding of Temporal Relations in Semantic Memory: Cognitive Load and Task-Evoked Pupillary Response*. Münster: Waxmann.
- Kubo, M. 2004. 日本製アニメとマンガの国際戦略. Research Institute of Economy, Trade and Industry (RIETI) BBL Seminar [Online]. Available from: <<http://www.rieti.go.jp/jp/events/bbl/04111601.html>> [Accessed 20 July 08]
- Kumar, R. 2005. *Research Methodology: A Step-by-Step Guide for Beginners*. 2nd ed. London: Sage Publications.
- Kusaka, K. and Koo, R. 2006. Time to polish the Japanese aesthetic sense. *Japan Echo*. 33 (2), pp 32 – 36.
- Lang, T. 2007. Documenting research in scientific articles: Guidelines for authors: 2. Reporting hypothesis tests. *Chest*. 131, pp 317 – 319.
- Lankshear, C. and Knobel, M. 2006. *New Literacies: Everyday Practices and Classroom Learning*. 2nd ed. Maidenhead: Open University Press.

- Lay, S. 2007. Audiences across the divide: Game to film adaptation and the case of Resident Evil. *Particip@tions* [Online]. 4(2). Available from: <http://www.participations.org/Volume%204/Issue%202/4_02_lay.htm> [Accessed 20 June 08]
- Ledoux, T. (ed.) 1997. *Anime Interviews: The First Five Years of Animerica, Anime and Manga Monthly (1992-97)*. San Francisco: Viz Communications.
- Lee, K.M. 2006a. *Building Good Measures* [Online]. California: University of Southern California. Available from: <http://www-rcf.usc.edu/~kwanminl/courses/comm301/lect4_1-Measurement%20validity%20and%20reliability.ppt> [Accessed 30 June 08]
- Lee, K.M. 2006b. *Validity in Research* [Online]. California: University of Southern California. Available from: <http://www-rcf.usc.edu/~kwanminl/courses/comm301/lect7_1-Validity%20and%20its%20threats.ppt> [Accessed 30 June 08]
- Leonard, S. 2004. *Progress Against the Law: Fan Distribution, Copyright, and the Explosive Growth of Japanese Animation* [Online] Version 1.11. Massachusetts: MIT. Available from: <<http://web.mit.edu/seantek/www/papers/progress-columns.pdf>> [Accessed 2 September 08]
- Leppihalme, R. 1997. *Culture Bumps: An Empirical Approach to the Translation of Allusions*. Clevedon: Multilingual Matters.
- Levi, A. 2006. The Americanization of anime and manga: Negotiating popular culture. *IN: Brown, S. (ed.) Cinema Anime*. New York: Palgrave Macmillan, pp 43-46.
- Lewis, P.E. 1981. Vers la traduction abusive. *Proceedings from the colloquium Les fins de l'homme 23 July – 2 August at Cerisy-la-Salle*. Paris: Galilée, pp 253-261.
- Lewis, P.E. 1985. The measure of translation effects. *IN: Graham, J.F. (ed.) Difference in Translation*. Ithaca: Cornell University Press, pp 31-62.
- Lovegren, S. 2003. Cartoons for grown-ups, Japan's anime draws millions. *National Geographic News* [Online]. Available from: <http://news.nationalgeographic.com/news/2003/09/0916_030916_milleniumpactress.html> [Accessed 15 July 06]
- Man, D.W.K. and Wong, M.S.L. 2007. Evaluation of computer-access solutions for students with quadriplegic athetoid cerebral palsy. *American Journal of Occupational Therapy*. 61, pp 355–364.

- Manzenreiter, W. 2007. The mangatzation of the world: Japanese popular culture, cultural diplomacy and the new international division of labour. *Japan Aktuell*. 4, pp 4-23.
- Martí Ferriol, J. L. 2007. An Empirical and descriptive study of the translation method for dubbing and subtitling. *Linguistica Anterpensia*. 6, pp 171-183.
- Mas López, J. 2004. From Tokyo to Barcelona, translating Japanese anime into Catalan. *The Globalisation Insider* [Online]. 9.
Available from:
<http://www.lisa.org/globalizationinsider/2004/09/from_tokyo_to_b.html>
[Accessed 29 September 08]
- Mayoral, R., Kelly, D., and Gallardo, N. 1988. Concept of constrained translation. Non-linguistic perspectives of translation. *Meta*. 33 (3), pp 356-367.
- McCarthy, H. 1999. *Hayao Miyazaki: Master of Japanese Animation: Films, Themes, Artistry*. Berkeley: Stone Bridge Press.
- McQuail, D. 1994. *Mass Communication Theory. An Introduction*. London: Sage.
- Miller, W.G. 2008. *OpenStat*. Ver. 1.9.
- Moran, S. 2008. *The Effect of Linguistic Variation on Subtitle Reception*. Audiovisual Translation: Multidisciplinary Approaches Conference, Montpellier, France, 19-20 June 2008.
- Musbach, G. 2004. Re: Question: The best anime DVD release. *news:rec.arts.anime.misc 17 November 2004* [Internet discussion list]. Available from:
<<http://rec.arts.anime.misc>>
[Accessed 17 November 04]
- Napier, S. 2001. *Anime from Akira to Princess Mononoke: Experiencing Contemporary Japanese Animation*. New York: Palgrave Macmillan.
- Napier, S. 2005. *Anime from Akira to Howl's Moving Castle: Experiencing Contemporary Japanese Animation*. New York: Palgrave Macmillan.
- Natsume, F. 2000. Japan's manga culture. *The Japan Foundation Newsletter*. 27(3/4), pp 1-6.
- Nesbit, J.C. and Hadwin, A.F. 2006. Methodological issues in educational psychology. IN: Alexander, P.A., and Winne, P.H. (eds.) *Handbook of Educational Psychology*. 2nd ed. New Jersey: Lawrence Earlbaum Associates and APA, pp 825-848.
- Nornes, A.M. 1999. For an abusive subtitling. *Film Quarterly*. 52 (3), pp 17-34.
- Nornes, A.M. 2004. For an abusive subtitling. IN: Venuti, L. (ed.) *The Translation Studies Reader*. 2nd ed. New York: Routledge, pp 447-469.

- Orero, P. 2008. Three different receptions of the same film: 'the pear stories project' applied to audio description. *European Journal of English Studies*. 12 (2), pp 179 -193.
- Orero, P. and Vercauteren, G. 2008. *The Audio Description of Facial Expressions*. Audiovisual Translation: Multidisciplinary Approaches Conference, Montpellier, France, 19 -20 June 2008.
- Ortabasi, M. 2006. Indexing the past: visual language and translatability in Kon Satoshi's Millennium Actress. *Perspectives: Studies in Translatology*. 14 (4), pp 278-291.
- O'Brien, S. 2006. Eye tracking and translation memory matches. *Perspectives - Studies in Translatology*. 14 (3), pp185-205.
- O'Brien, S. 2008. Processing fuzzy matches in translation memory tools: an eye-tracking analysis. *IN: Göpferich, S., Jakobsen, A.L. and Mees, I.M. (eds.) Looking at Eyes. Eye-Tracking Studies of Reading and Translation Processing. Copenhagen Studies in Language*, pp 79-102.
- O'Donnell, R.D. and Eggemeier, F.T. 1986. Workload assessment methodology. *IN: Boff, K.R., Kaufman, L., and Thomas, J.P. (eds.) Handbook of Perception and Human Performance: Vol. II Cognitive Processes and Performance*. New York: Wiley Interscience, pp 42/1-42/9.
- O'Hagan, M. 2006a. Manga, anime and videogames: Globalising Japanese cultural production. *Perspectives: Studies in Translatology*. 14 (4), pp 242-247.
- O'Hagan, M. 2006b. *Approaching Manga, Anime and Videogames from a Translation Studies Perspective*. The International Conference on Asian Comics, Animation and Gaming, Toronto, Canada 18-19 May 2006.
- O'Hagan, M. 2007. Impact of DVD on translation: Language options as an essential add-on feature. *Convergence: The International Journal of Research into New Media Technologies*. 13 (2), pp 157-168.
- O'Hagan, M. 2008a. Fan translation networks: An accidental translator training environment? *IN: Kearns, J. (ed.) Translator and Interpreter Training: Methods and Debates*. London: Continuum, pp 158-183.
- O'Hagan, M. 2008b. *Manga, Anime and Videogames: An Interdisciplinary Initiative in Japanese, Multimedia and Translation Studies*. Learning Innovation Unit Seminar, Dublin City University, Dublin 13 March 2008.
- Paas, F., Tuovinen, J.E., Tabbers, H., and Van Gerven, P.W.M. 2003. Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*. 38 (1), pp 63-71.
- Pedersen, J. 2005. *How is Culture Rendered in Subtitles?* EU High Level Conference Series – Challenges of Multidimensional Translation, Saarbrücken, Germany, 2-6 May 2005.

- Pedersen, J. 2007. Cultural interchangeability: The effects of substituting cultural references in subtitling. *Perspectives: Studies in Translatology*. 15 (1), pp 30 – 48.
- Perego, E. and Del Missier, F. 2008. *Is a Reading Situation Better than Another for Subtitled Film Viewers?* Audiovisual Translation: Multidisciplinary Approaches Conference, Montpellier, France, 19 -20 June 2008.
- Pérez González, L. 2006. Fansubbing anime: Insights into the 'butterfly effect' of globalisation on audiovisual translation. *Perspectives: Studies in Translatology*. 14 (4), pp 260-277.
- Pérez González, L. 2007. Intervention in new amateur subtitling cultures: a multimodal account. *Linguistica Anterpensia*. 6, pp 67 – 79.
- Pettit, Z. 2004. The audio-visual text: Subtitling and dubbing different genres. *Meta*. 49 (1), pp 25–38.
- Pettit, Z. 2007. Translating verbal and visual language in *The Piano*. *Perspectives: Studies in Translatology*. 15 (3), pp 177-190.
- Phillips, S. 1996. Manga für ein deutschsprachiges Publikum: Möglichkeiten der Übertragung von Text-Bild-Verbindungen. *Japanstudien*. 8, pp 193-210.
- Posner, M.I. 1980. Orienting of attention. *Quarterly Journal of Experimental Psychology*. 32, pp 3-25.
- Praet, C., Verfaillie, K., De Graef, P., Van Rensbergen, J., and d'Ydewalle, G. 1990. A one-line text is not half a two-line text. *IN: Groner, R., d'Ydewalle, G., and Parham, R. (eds.) From Eye to Mind: Information Acquisition in Perception, Search, and Reading*. Amsterdam: North-Holland, pp 205-213.
- Price, S. 2001. Cartoons from another planet: Japanese animation as cross-cultural communication. *Journal of American and Comparative Cultures*. 24(1/2), pp 153-169.
- Rayner, K. 1992. Eye movements and visual cognition: Introduction. *IN: Rayner, K. (ed.) Eye Movements and Visual Cognition: Scene Perception and Reading*. Berlin: Springer Verlag, pp 1-7.
- Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*. 124 (3), pp 372-422.
- Richards, I.A. 1929. *Practical Criticism. A Study of Literary Judgement*. New York: Harcourt, Brace and Co.
- Richie, D. 1991. Donald Richie on subtitling Japanese films. *Mangajin*. 10, p 16.
- Rovira Esteva, S. 2007. Translating Chinese pop fiction. *Perspectives: Studies in Translatology*. 15 (1), pp 15-29.

- Ruch, W. 1992. Assessment of appreciation of humor: Studies with the 3WD humor test. *IN: Spielberger, C.D., and Butcher, J.N. (eds.) Advances in Personality Assessment, 9.* New Jersey: Lawrence Erlbaum Associates, pp 27-75.
- Salkind, N.J. 2008. *Statistics for People Who (Think They) Hate Statistics.* 3rd ed. California: Sage Publications.
- Salojärvi, J., Puolamäki, K., Simola, J., Kovanen, L., Kojo, I., and Kaski, S. 2005. Inferring relevance from eye movements: Feature extraction (Technical Report A82). *Helsinki University of Technology, Publications in Computer and Information Science* [Online]. Available from: <<http://www.cis.hut.fi/eyechallenge2005/>> [Accessed 9 April 08]
- Santos, J. 2007. Data on Data: Viewer responses to Star Trek: The Next Generation. *Particip@tions* [Online]. 4(1). Available from: <http://www.participations.org/Volume%204/Issue%201/4_01_santos.htm> [Accessed 12 June 08]
- Schilling, M. 1997. *The Encyclopaedia of Japanese Pop Culture.* New York: Weatherhill.
- Schodt, F. 2007. *The Astro Boy Essays: Osamu Tezuka, Mighty Atom, and the Manga/Anime Revolution.* Berkeley: Stone Bridge Press.
- Schultheis, H. 2004. *Pupillengröße und Kognitive Belastung.* Diplomarbeit. Saarbrücken Universität.
- Schultheis, H. and Jameson, A. 2004. Assessing cognitive load in adaptive hypermedia systems: Physiological and behavioural methods. *IN: Neijdl, W., and de Bra, P. (eds.) Adaptive Hypermedia and Adaptive Web-based Systems.* Eindhoven: Springer Verlag, pp 18-24.
- Schultz, N. 2008. Fan fiction. *Britannica Online Encyclopedia* [Online]. Available from: <<http://www.britannica.com/EBchecked/topic/1017613/fan-fiction>> [Accessed 25 August 08]
- Schlingensiepen, K.H., Campbell, E.W., Legge, G.E., and Walker, T.D. 1986. The importance of eye movements in the analysis of simple patterns. *Vision Research.* 26, pp 1111-1117.
- Shapiro, S.S. and Wilk, M.B. 1965. An analysis of variance test for normality (complete samples). *Biometrika.* 52 (3/4), pp 591 – 611.
- Shūji, T. 2006. Cultural assets and liabilities. *Japan Echo.* 4, pp 26–27.
- Sperber, D. and Wilson, D. 1986. *Relevance: Communication and Cognition.* 1st ed. Oxford: Blackwell.

- Sperber, D. and Wilson, D. 1995. *Relevance: Communication and Cognition*. 2nd ed. Oxford: Blackwell.
- Steinhauer, S.R. 2002. *Pupillary Responses, Cognitive Psychophysiology and Psychopathology* [Online]. Pittsburgh: University of Pittsburgh School of Medicine. Available from:
<<http://www.wpic.pitt.edu/research/biometrics/Publications/PupilWeb.htm>>
[Accessed 22 August 08]
- Szarkowska, A. 2007. *Why are Some Forms of Address not Omitted in Subtitling? A Study Based on Three Selected Polish Soap Operas Broadcast on TV Polonia*. Intermedia Conference, Łódź, Poland, 13-14 April 2007.
- Taylor, C. 2003. Multimodal transcription in the analysis, translation and subtitling of Italian films. *The Translator*. 9 (2), pp 191-205.
- Titford, C. 1982. Sub-titling: Constrained translation. *Lebende Sprachen*. 27 (3), pp 113-116.
- Toury, G. 1991. Experimentation in translation studies: Prospects and some pitfalls. IN: Tirkonnen-Condit, S. (ed.) *Empirical Research on Translation and Intercultural Studies*. Tübingen: Narr, pp 55-66.
- TV Tokyo Corporation. 2006. *2006 Annual Report*. Tokyo: TV Tokyo Corporation.
- Valentini, C. 2006. *Mediating the Visual Through the Aural. Some Reflections on the Localisation of Films*. Languages and the Media Conference, Berlin, Germany, 26-27 October 2006.
- Van Rensbergen, J., d'Ydewalle, G., Eraerts, N., and Robben, G. 1986. A developmental perspective on reading subtitles: Television watching behaviour of children. *Bulletin of the British Psychological Society*. 39: A94.
- Venuti, L. 1995. *The Translator's Invisibility*. London: Routledge.
- Venuti, L. 2006. How to read a translation. *Words without Borders – The Online Magazine for International Literature* [Online]. Available from:
<<http://www.wordswithoutborders.org/article.php?lab=HowTo>>
[Accessed 21 August 06]
- Verney, S., Granholm, E. and Dionisio, D. P. 2001. Pupillary responses and processing resources on the visual backward masking task. *Psychophysiology*. 38, pp 76 – 83.
- White, C.W. 1976. Visual masking during pursuit eye movements. *Journal of Experimental Psychology: Human Perception and Performance*. 2, pp 469-478.
- Widler, B. 2004. A questionnaire among audiences of subtitled films in Viennese cinemas. *Meta*. 49 (1), pp 98-101.

- Wierzbicka, A. 1995. Kisses, bows and handshakes. *Semiotica*. 103 (3/4), pp 207–52.
- Wilson, D. and Matsui, T. 1998. Recent approaches to bridging: Truth, coherence and relevance. *UCL Working Papers in Linguistics*. 10, pp 73-200.
- Wilson, D. and Sperber, D. 2002. Relevance theory. *UCL Working Papers in Linguistics*. 14, pp 249-287.
- Wilson, G.F., and Eggemeier, F.T. 1991. Psychophysiological assessment of workload in multi-task environments. *IN: Damos, D.L. (ed.) Multiple-Task Performance*. London: Taylor and Francis, pp 207-216.
- Yatabe, K. 2006. *Saccadic Suppression During Reading Activity: Is the Spillover Effect Weaker after a Longer Saccade?* M.Sc. Thesis. University of Edinburgh.
- Yoshioka, H. 2007. *Japanization of Imagination? Globalization of Culture and Transformed Otherness*. „Cool Japan“: Bildexporte, Globalisierung, interkulturelle Begegnungen Mini-Symposium, Berlin, Germany, 31 October 2008.
- Yus, F. 2008. Alterations of relevance in cyber-media. *Universitas Psychologica*. 7 (3), pp 623-636.
- Zahlten, A. 2007. *Von Skandalen, Desinteresse und Zauberwörtern: Zum jüngsten Wandel von Filmkultur in Japan und seinen Auswirkungen in Europa*. „Cool Japan“: Bildexporte, Globalisierung, interkulturelle Begegnungen Mini-Symposium, Berlin, Germany, 31 October 2008.

Media

- Azumanga Daioh – Entrance (vol. 1)* 2004. [DVD]. Directed by Hiroshi Nishikiori. USA: ADV Films.
- Batman Returns* 1992. [35 mm]. Directed by Tim Burton. USA: Warner Bros. Pictures.
- Bridget Jones: The Edge of Reason* 2004. [DVD]. Directed by Beeban Kidron. USA: Working Title Films.
- Doki Doki School Hours: 1st hour* 2005. [DVD]. Directed by Yoshiaki Iwasaki. USA: Geneon Entertainment Inc.
- Duck Soup* 1933. [35 mm]. Directed by Leo Mc Carey. USA: Paramount Pictures.
- Millennium Actress* 2001. [35 mm]. Directed by Satoshi Kon. Japan: Bandai.
- Paniponi Dash! Lethal Lesson* 2007. [DVD]. Directed by Shin Onuma. USA: ADV Films.
- The Living Daylights* 1987. [35 mm]. Directed by John Glen. USA: MGM.
- The Matrix: Special Edition* 1999. [DVD] Directed by Andy and Larry Wachowski. USA: Groucho II Film Partnership.

Appendix A: Preparation of stimuli for use with ClearView – DVD conversion, scene definition and areas of interest (AOIs)

DVD conversion

The original DVD was first decrypted on the PC by using the freely available software, *DVD Shrink*, which was necessary to bypass security measures in the DVD recorder. After this, the newly decrypted DVD was put in a DVD player, which was hooked up to a separate DVD recorder by means of a SCART lead. The desired excerpts were then recorded to a DVD in both conditions: with subtitles and pop-up gloss, and with just subtitles. This DVD with the excerpts was put into the PC and the excerpts were copied onto PC before being converted to an .avi format file using the freely available software, *FlaskMPEG*.

Scene definition

The *scene definition* process divides the AV stimulus into segments, which can be analysed independently of the rest of the stimulus. Firstly, the define scenes tool must be selected in the analysis part of ClearView. As Figure A.1 illustrates, the user is presented with a *video window* displaying the stimulus (a), a colour coded *timeline* for the stimulus (b), a *list of defined scenes* (c), a *new scene group button* (d) and a *scene group details area*, where a still image representation of the selected scene can be seen, along with its representative colour on the timeline (e).

Using this window, the timeline is used to scan to the segment of the stimulus to be distinguished and a marker is set at the beginning of the area by pressing ctrl+m. The timeline is then used to scan to the end of the segment where ctrl+m is used again to set an ending marker. After doing this, the New Scene Group button is clicked on, and the Define New Scene Group window appears. Here, the scene group selected is given a name and a still image graphic representation for the image is displayed. This graphic representation is determined by the image in the video window, which can be changed by dragging the slider in the timeline.

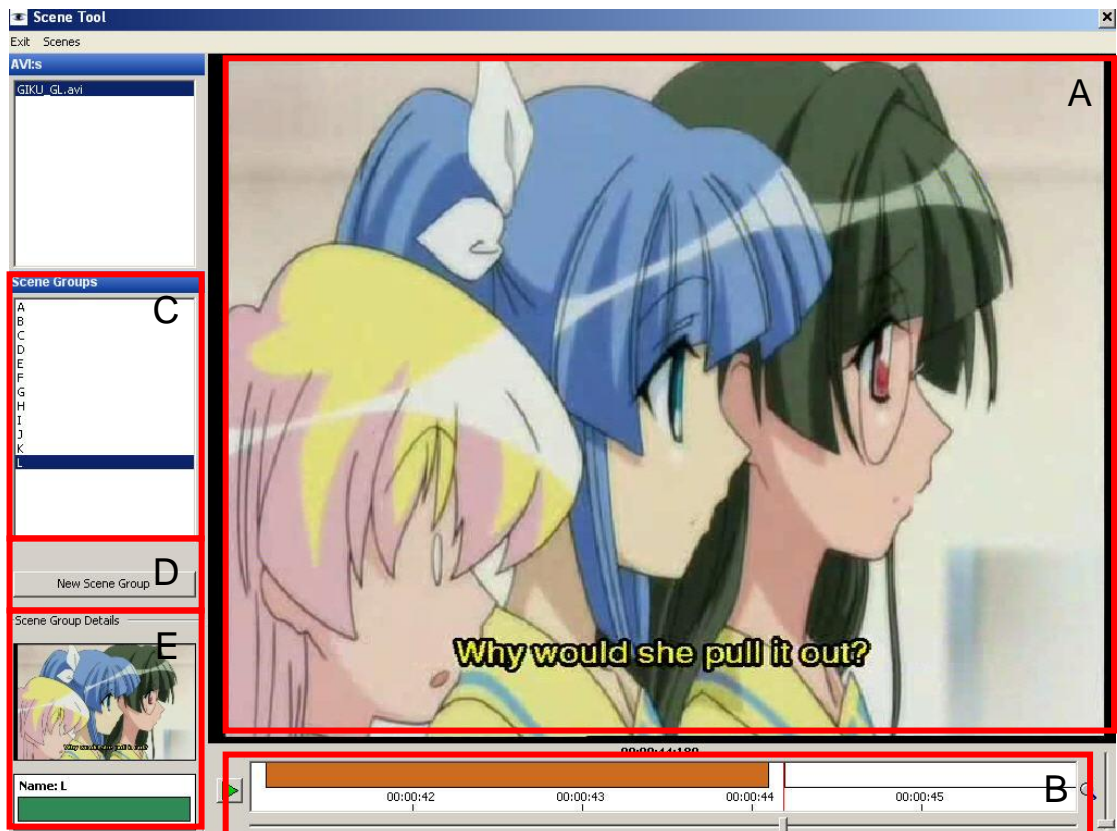


Figure A.1. Scene definition tool

After giving the segment a name, the name appears in the list of defined scenes. The name is then clicked on in the list, and the graphic representation appears in the scene group details area, and below that, the representative colour for the segment. From here, the user must drag and drop the colour from the lower left corner of the screen onto the area of the timeline where the segment they have just defined is, colouring the appropriate area of the timeline (see Figure A.2).

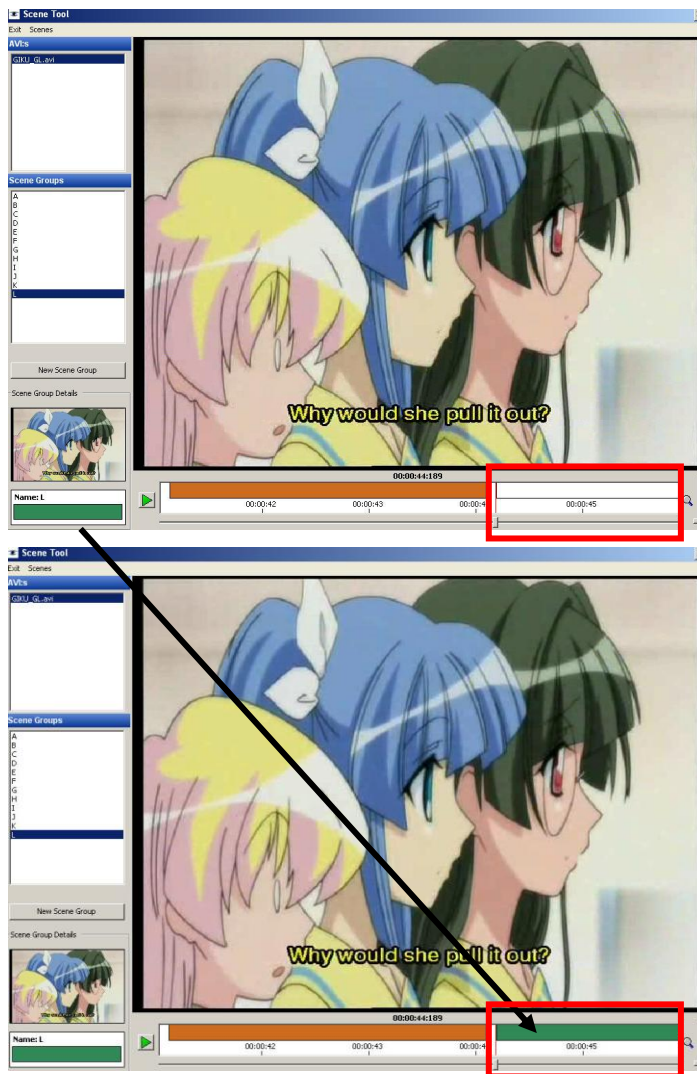


Figure A.2. Colouring segment during scene definition

When all of the scenes for analysis in the study are defined, the next step is to go one layer deeper in the segmentation and to divide each of these segments into areas of interest (AOIs).

Area of interest definition

To define AOIs, the AOI definition tool is used. The AOI definition screen has the following content, as illustrated in Figure A.3: a *list of all the scenes defined* from the scene definition process outlined above (a), a *current frame window*, which displays the static graphic representation of the segment selected (b), the *AOI insertion toolbar* (c) and the *AOI list* for the active segment (d).

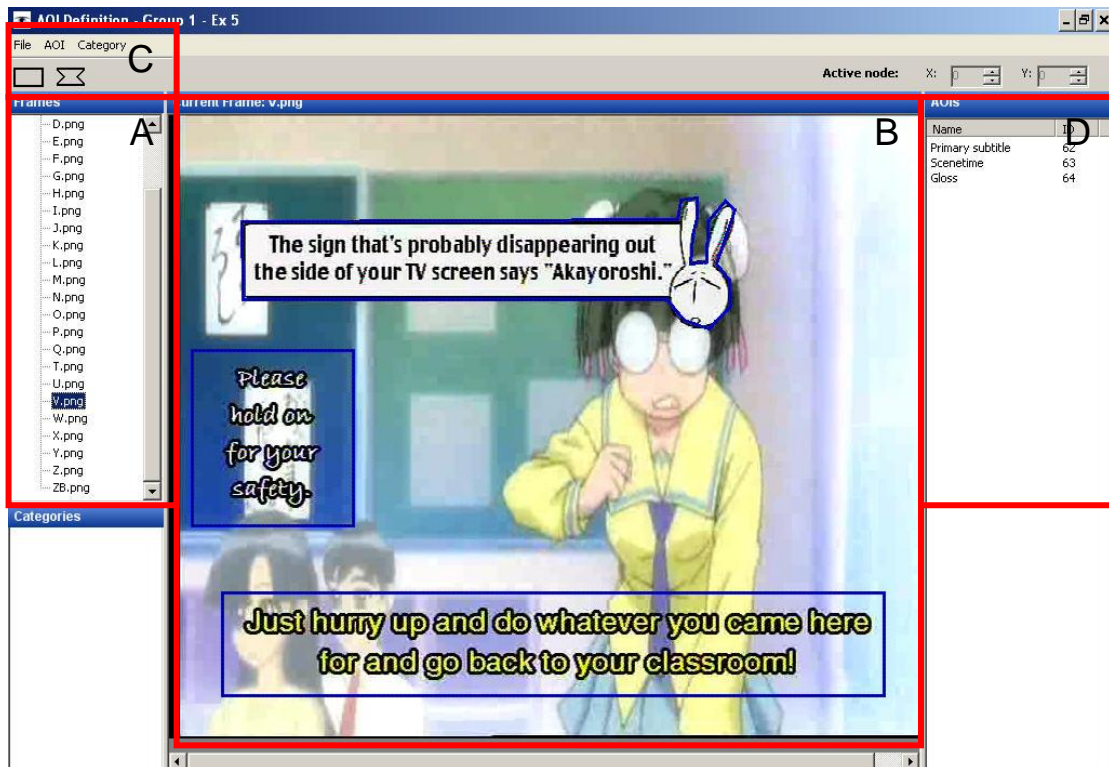


Figure A.3: AOI definition screen

To add an AOI, the segment is selected from the list of predefined scenes. Upon selection of the segment, the static representation of that segment appears in the current frame window. Here, the AOI insertion toolbar is used to select either a quadrilateral or free-form AOI. After selecting the appropriate shape, the area of interest is highlighted on the current frame window and given a name. This is carried out for all scenes. AOI templates may be created and reused if the AOI appears in the same place in several scenes.

Appendix B: Pilot study questionnaire

Anime Perception Study Survey

This survey is being carried out to investigate the effects that subtitles have on viewers' perceptions of culturally and genre specific nonverbal cues in an anime series made for TV broadcast.

Please note that some of the scenes may cause mild offence

.....

- Where there are multiple-choice answers, please circle the appropriate one.
- When asked to rate your understanding, please rate your understanding on a scale of 0 to 6, with 0 meaning didn't understand at all, and 6 meaning fully understood.

SECTION I

1. Age:
2. Gender: M F
3. Japanese language ability (0 – none 6 – native)
0 1 2 3 4 5 6
4. English language ability (0 – none 6 – native)
0 1 2 3 4 5 6
5. What is anime's country of origin?
Japan China Korea Vietnam USA France
6. Define anime.

-
7. Rate your level of interest in anime.
Low Medium High
 8. Do you consider yourself an anime fan?
Yes No

SECTION II

Kinesic

CLIP 1:

1. Do you understand the bowing exchange between the two men we see towards the end of the clip?
Please rate your understanding.
Not at all 0 1 2 3 4 5 6 *Fully*
2. Why do the men lower their bows in succession?

Clip 2:

3. Do you understand the significance of the young man reading the paper, picking his ear and slapping the wet towel towards the end of the clip?

Please rate your understanding.

Not at all 0 1 2 3 4 5 6 *Fully*

4. What is the significance of the man doing all these things?
-

Cultural signsClip 1:

5. Do you know what game the boy is playing here?

Please rate your understanding.

Not at all 0 1 2 3 4 5 6 *Fully*

6. What does the boy want to do to decide the seating arrangement?
-

Clip 2:

7. Do you understand the reference being made to the prawn and fish?

Please rate your understanding.

Not at all 0 1 2 3 4 5 6 *Fully*

8. What does the reference to the prawn and the fish mean?
-

Genre specific signsClip 1:

9. Do you understand what the black symbol on the girl's cheek means?

Please rate your understanding.

Not at all 0 1 2 3 4 5 6 *Fully*

10. What does the black symbol on the girl's cheek symbolise?
-

Clip 2:

11. Do you understand why the boy had a nosebleed?

Please rate your understanding.

Not at all 0 1 2 3 4 5 6 *Fully*

12. What does the nosebleed symbolise?
-

SECTION III

1. Did you find any of the clips disturbing or unusual in any way? Explain.
-

2. Which strategy do you think subtitlers should take when they encounter a culturally specific or genre specific image?

- A – Ignore it – allow the viewer to guess its meaning.
 - B – Explain it in the subtitle.
 - C – Explain it in a note at the top of the screen.
 - D – Explain it in a translator’s notes packaged with the product.
 - E – Change the meaning of the speech to suit the image.
 - F – Other:
-

3. When you encounter a culturally specific image you don’t understand in a subtitled film, do you:

- A – Ignore it and assume if it isn’t subtitled it isn’t necessary.
 - B – Try to find its meaning out on your own (e.g. friends, internet)
 - C – Guess its meaning by the context (e.g. dialogue, other images)
 - D – Other:
-

4. Did you find the subtitle speeds too fast?
Yes No Sometimes

5. Did you find the subtitle speeds too slow?
Yes No Sometimes

6. If you answered yes to question 4 or 5, please explain when you found this.
-

If you have any other comments, please write them here.

Appendix C: Mean values for fixation-based and pupillometric dependent variables

Percentage of subtitles skipped

Experiment condition	1 line gloss	2 line gloss	1 line solo	2 line solo
Subject #				
1	8.84354	4.34783	33.33333	12.5
3	1.36054	0	8.33333	0
4	0.68027	0	12.5	0
6	0	0	12.5	12.5
7	0.68027	0	12.5	12.5
8	1.36054	0	25	12.5
10	0.68027	0	41.66667	12.5
11	30.8642	0	44.44444	0
12	3.7037	0	33.33333	0
13	3.7037	0	33.33333	0
14	6.17284	0	44.44444	0
16	3.08642	0	33.33333	0
17	14.19753	0	44.44444	0
19	1.85185	0	33.33333	0

Percentage of gaze time spent in subtitle area

Experiment condition	1 line gloss	2 line gloss	1 line solo	2 line solo
Subject #				
1	42.26308	38.86667	14.80074	22.26409
3	62.50417	74.99822	39.93609	61.09343
4	55.42373	68.15079	28.53977	56.10987
6	50.84222	73.67672	36.20801	53.28789
7	36.5597	46.94985	26.20205	31.43229
8	56.76731	69.90914	35.14743	54.33167
10	36.50475	52.29346	14.30612	36.10523
11	16.69951	40.72932	17.03569	22.67215
12	34.40314	51.18681	17.55353	32.10356
13	33.86978	45.19621	17.29444	31.30633
14	28.13789	39.14309	21.25527	39.68838
16	47.97642	64.58257	23.00136	41.193
17	36.52897	55.63922	14.69245	47.09742
19	49.31294	62.99605	21.93501	24.26007

Mean fixation duration (ms)

Experiment condition	1 line gloss	2 line gloss	1 line solo	2 line solo
<i>Subject #</i>				
1	164.7985	172.2571	264.6353	182.6326
3	204.0777	218.2594	281.6874	232.5523
4	211.8171	205.6279	277.967	218.2407
6	233.5103	239.4036	321.3282	223.6583
7	146.6863	150.6443	182.8122	170.3151
8	216.491	202.1568	251.7493	215.8739
10	166.6595	172.1927	188.4358	177.867
11	186.25	189.25	191.3723	200.9325
12	157.9467	165.7	184.476	173.3978
13	174.2167	172.6667	187.9313	175.4139
14	212.0833	196.3571	244.4802	226.1099
16	234.3056	283.1	273.4038	221.3309
17	192.7733	167	222.8732	182.8537
19	184.1833	231.2	245.3749	218.7899

Word fixation probability

Experiment condition	1 line gloss	2 line gloss	1 line solo	2 line solo
<i>Subject #</i>				
1	0.560072	0.503824	0.997015	0.974863
3	0.925015	0.752683	1.316418	1.285733
4	0.774453	0.729956	1.123413	1.093531
6	0.811954	0.739917	0.941046	0.936307
7	0.930769	0.709361	1.093444	1.068804
8	0.987489	0.900433	1.281095	1.269766
10	0.708276	0.693903	1.064876	1.045792
11	0.577381	0.339286	0.590619	0.598039
12	0.774762	0.470238	1.01104	1.002042
13	0.632381	0.464286	0.95598	0.976722
14	0.830952	0.541667	0.63416	0.636367
16	0.920635	0.428571	0.982767	0.98774
17	0.72619	0.821429	1.004686	1.010124
19	0.690476	0.386905	1.146414	1.162421

Median pupil size (mm)

Experiment condition	1 line gloss	2 line gloss	1 line solo	2 line solo
<i>Subject #</i>				
1	3.643916	3.667196	3.637104	3.729438
3	4.483108	4.571478	4.4715	4.552188
4	3.907861	3.914435	3.874333	3.985
6	4.396267	4.410283	4.249375	4.353813
7	4.211101	4.28763	4.231833	4.2865
8	3.791804	3.815522	3.753292	3.744375
11	4.433015	4.44969	4.399313	4.71325
12	5.126281	5.135328	5.0535	5.083
13	3.509923	3.553293	3.594125	3.82275
14	3.826895	3.88769	3.880313	3.85275
16	5.302778	5.305069	5.377438	5.4805
17	3.988775	4.008931	3.940938	4.0005
19	4.646824	4.766638	4.768625	4.96575

Appendix D: Email from *Tobii* re: screen resolution

From: "Joakim Isaksson" <support@tobii.com>
To: colm.caffrey@dcu.ie
Cc:
Subject: [Tobii #DAQ-244723]: Questions re: Tobii
Date: Mon, 01 Sep 2008 14:10:55 +0200

Hello Colm,

The quality of the eye tracking data itself will not be affected by the screen resolution used for recording. The only affect it has is that the lower resolution will make the details in the stimulus less clear, the gaze data will still be the same however.

Best regards,

Joakim Isaksson

Joakim Isaksson
Support Engineer

Support: +46 (0)8 522 950 10
Direct: +46 (0)8 522 951 84
Mobile: +46 (0)70 916 16 44
Fax: +46 (0)8 30 14 00
E-mail: joakim.isaksson@tobii.com
Web: www.tobii.com

=====
colm.caffrey@dcu.ie (Client) Posted On: 27 Aug 2008 11:18

Appendix E: Content questionnaire from experiment

Please mark the appropriate box and answer ALL questions. If there is any question you do not understand or if you have any questions, please ask me.

Excerpt 1

1. Do you know what the sequence with the gun sight following the girl represents?

Not at all 1 2 3 4 5 6 Completely

2. What movie does this sequence make a reference to?
3. How did you find the speed of the main dialogue subtitle in this excerpt?
Too slow 1 2 3 4 5 Too fast
4. What does the girl say she will reveal about Rebecca Miyamoto?

Excerpt 2

1. What will the dark-haired girl give Becky if she goes to her house?
2. Do you understand the emoticon (°Д°) in this clip?
Not at all 1 2 3 4 5 6 Completely
3. What does the emoticon (°Д°) depict?
4. How did you find the speed of the main dialogue subtitle in this excerpt?
Too slow 1 2 3 4 5 Too fast

Excerpt 3

1. What does Becky ask the plain girl to do for her?
2. Do you understand the reference of the eye in the palm of the hand?
Not at all 1 2 3 4 5 6 Completely
3. What is the eye in the palm of the hand a reference to?
4. How did you find the speed of the main dialogue subtitle in this excerpt?
Too slow 1 2 3 4 5 Too fast

Excerpt 4

1. What does the girl want Becky to say to prove she has presence?

2. Do you understand why we see the fruits near the end of this clip?
 Not at all 1 2 3 4 5 6 Completely
3. Why do we see peaches, persimmons and chestnuts falling after Becky guesses the wrong name?
4. How did you find the speed of the main dialogue subtitle in this excerpt?
 Too slow 1 2 3 4 5 Too fast

Excerpt 5

1. What does the girl tell the old teacher he'd better not pull out a picture of?
2. Do you know what the photo frame surrounded by the black ribbon represents?
 Not at all 1 2 3 4 5 6 Completely
3. What is a photo frame with a black ribbon used for in Japan?
4. How did you find the speed of the main dialogue subtitle in this excerpt?
 Too slow 1 2 3 4 5 Too fast

Excerpt 6

1. Why does the rabbit suggest the drinks may be warm?
2. Do you know what the white doll is?
 Not at all 1 2 3 4 5 6 Completely
3. What was this type of doll originally designed for?
4. How did you find the speed of the main dialogue subtitle in this excerpt?
 Too slow 1 2 3 4 5 Too fast

Excerpt 7

1. Becky thinks worrying about unnecessary stuff shows that you have what sort of problems?
2. Do you know what the picture of the volcano parodies?
 Not at all 1 2 3 4 5 6 Completely

3. The picture of the volcano parodies the work of which Japanese artist?
4. How did you find the speed of the main dialogue subtitle in this excerpt?

Too slow 1 2 3 4 5 Too fast

Excerpt 8

1. What sort of work does the girl in the pink dress say that she does not do?
2. Do you know what the green creature in the water is?

Not at all 1 2 3 4 5 6 Completely

3. What is a kappa?
4. How did you find the speed of the main dialogue subtitle in this excerpt?

Too slow 1 2 3 4 5 Too fast

Excerpt 9

1. What do the girls say is the most important for a typical reporter?
2. Do you understand what the figure in white with the lowered head and extended hand represents?

Not at all 1 2 3 4 5 6 Completely

3. Who is a white kimono traditionally meant for?
4. How did you find the speed of the main dialogue subtitle in this excerpt?

Too slow 1 2 3 4 5 Too fast

Excerpt 10

1. Do you understand what the character with the white speech bubble and red writing parodies?

Not at all 1 2 3 4 5 6 Completely

2. What videogame does this scene make a reference to?
3. How did you find the speed of the main dialogue subtitle in this excerpt?

Too slow 1 2 3 4 5 Too fast

4. Why does Ichijo not want one of the toys?

Excerpt 11

1. What sort of fairy is born from the rabbit's head?
2. Do you know what the scene with the cats and the girl parodies?
Not at all 1 2 3 4 5 6 Completely
3. What character does this sequence make a reference to?
4. How did you find the speed of the main dialogue subtitle in this excerpt?
Too slow 1 2 3 4 5 Too fast

General questions:

1. Which option do you think is the best for dealing with cultural items in translation on DVD?
 Explain it on screen
 Include it in subtitle
 Explain it in a booklet included with DVD
 Ignore it
 Change the dialogue to fit the image
 Other: _____
2. How did you find the pop-up notes?
3. Would you watch an entire DVD subtitled in this way?
Yes No
4. Did you find anything disturbing about the content of the excerpts?
Yes No
5. If so, please elaborate:

Appendix F: Recruitment email

Dear students,

I will be carrying out a study into the reception of anime in the month of March. I am looking for people to participate in the study. It will involve filling out a short questionnaire, then coming to the Henry Grattan and watching some short anime clips on an eye tracking monitor and answering some questions on the clips. It should take no longer than one hour. Please note that participation in the study is not possible if you:

- have any condition which adversely affects your vision or reading.
- are or ever have been a substance abuser.
- have diabetes.
- have ever had a brain or serious head injury.

If you are interested in taking part, please reply to this email,

Best regards,
Colm Caffrey

Appendix G: Email questionnaire

SURVEY OF ANIME PERCEPTION

Please complete the enclosed questionnaire and return it by E-mail to colm.caffrey@dcu.ie. If you have any problems or queries, please enquire to the same address.

1. Instructions for Respondents

1. Please select your answer/answers from the set of answers provided below each question by placing a check mark in the appropriate checkbox/checkboxes (point the mouse on the shaded checkbox and click once with the left button of the mouse).
2. Use the textbox below the answer "Other (please specify)" to type in answers not contained in the provided set (point the mouse on the shaded textbox, click once with the left button of the mouse [the textbox is highlighted], and type in your answer).
3. A number of questions ask you to reply in free form text, by typing in the textbox placed below the question (point the mouse on the shaded textbox, click once with the left button of the mouse [the textbox is highlighted], and type in your answer).

Note: When you see the word "content" in questions it refers to all audiovisual media. This includes, but is not limited to: film, TV programmes, cartoons, DVD, VCR and VCD.

General details

1 Gender:

Female.....

Male.....

2 Age:

3 Are you a native English speaker?

Yes.....

No.....

4 Can you understand Japanese?

Yes.....

No.....

2. Viewing preferences and familiarity

Content

5 How often do you watch content on PC?

- Several times per week.....
- Weekly.....
- Several times per month.....
- Monthly.....
- Several times per year.....
- Less often.....
- Never.....

6 How often do you watch subtitled content?

- Several times per week.....
- Weekly.....
- Several times per month.....
- Monthly.....
- Several times per year.....
- Less often.....
- Never.....

7 How often do you watch subtitled content on PC?

- Several times per week.....
- Weekly.....
- Several times per month.....
- Monthly.....
- Several times per year.....
- Less often.....
- Never.....

8 Do you prefer content to be dubbed or subtitled?

- Dubbed.....
- Subtitled.....
- No preference.....

Anime and manga

9 Do you like manga?

- Yes.....
- No.....
- No opinion on it.....
- Don't know what it is.....

10 How often do you read manga?

- Several times per week.....
- Weekly.....
- Several times per month.....
- Monthly.....
- Several times per year.....
- Less often.....
- Never.....

11 Do you like anime?

- Yes.....
- No.....
- No opinion on it.....
- Don't know what it is.....

12 How often do you watch anime?

- Several times per week.....
- Weekly.....
- Several times per month.....
- Monthly.....
- Several times per year.....
- Less often.....
- Never.....

13 Are you familiar with the anime series Paniponi Dash?

- Yes.....
- No.....