Reception Studies and Audiovisual Translation: Eye Tracking Research at the Service of Training in Subtitling

Elena Di Giovanni

University of Macerata, Italy

We are in an era of
prestige television,
with unprecedented
choice and quality.
(*The Guardian*, 21 August, 2019)

Abstract

In a recent article about viewers' choice of shows and series, The Guardian *referred to contemporary television as "comfort TV" and "prestige television", including all those streaming services that have recently pushed the very concept of television beyond traditional borders, while also giving it a new, vibrant life. 'Comfort' is most likely referred to the ease of consumption, both in terms of space (with portable technologies) and time (no longer univocally fixed by broadcasters). 'Prestige', as the quote says, is related to the incredible, and increasing, array of choices available; whereas quality can be referred to overall production standards, but perhaps not always to translation.*

As consumption of contemporary television, including streaming services, increases and diversifies, the need to understand, monitor and cater for the viewers' needs and expectations is ever more important, also through the lens of translation. This article discusses a recent experiment on the reception of a subtitled TV series by young Italian viewers, carried out using eye tracking and questionnaires. The experiment is here presented mainly to highlight the relevance, and the great potential, of reception studies applied to the training of audiovisual translators, and more specifically of subtitlers. After all, learning an eminently practical skill has to encompass as thorough a knowledge of its users as possible. Inspired by the curiosity of a subtitling company about viewers' comprehension and appreciation of

subtitled audiovisual texts viewed on screens of different sizes, this experiment was conducted using authentic subtitled materials and, as often happens with experimental research, it provided results that were largely unexpected, also bringing to the fore issues (such as reading speed) that had not been considered central at the onset.
In the following sections, key concepts in which the experiment is grounded will be explored, followed by the discussion of issues which have emerged from the experiment itself. As for the results, only a few will be presented, with a view to highlighting their relevance for the training of subtitlers.

1. **Audiovisual translation and audience research**

Within the realm of audiovisual translation, audience research has been booming over the past few years, pushing the field of study beyond a rather stale descriptivism. As recently as 2009, Yves Gambier observed that "very few studies have dealt with the issue of reception in AVT, and even fewer have looked at it empirically, even though we continually make references to readers, viewers, customers" (2009: 52). Ten years on, a fairly conspicuous number of articles, several handbook entries and a few dedicated books bear witness to a strong move towards a long-due, hopefully ever more systematic consideration of the audience needs and preferences.

Looking back, one clear reason for the slow inception of reception studies in AVT, in itself a relatively young field of research, is certainly related to modes of consumption, with viewers largely -and increasingly- consuming audiovisual texts in private settings, with great variation according to age, occupation, education, geographical factors, etc. Nevertheless, understanding and catering for audience reception of translated audiovisual texts is as important as it is for original productions, especially if we bear in mind that, on average, 55 to 65% of foreign grosses for blockbuster films come from translated versions (http://www.boxofficemojo.com).[1]

As studies of perception and/or reception (Di Giovanni, 2018) are on the increase in AVT, several pathways have come to be increasingly defined. Corpus-based studies, which have been developed for some

[1] See, for instance, the figures for recent blockbusters such as *Avengers: Endgame* or *Joker*, both released in 2019. Revenues from foreign distribution amount to 69.3% and 68.5% respectively. Considering that foreign distribution also includes some English-speaking countries, and that figures are not yet stable for these films, we could say that 62 to 65% of the revenues for these films actually come from translated versions (figures checked on 10 December 2019).

time (see Pavesi, 2013, 2014, 2019), have had the great merit of objectively depicting instances of standardization or lack of fluency in translation, manifestations of dubbese (synchronically and across the decades) recurrent register shifts between specific language pairs, and so on. Questionnaires, administered at specific screenings or delivered online (Di Giovanni, 2016, Romero Fresco, 2015, Caffrey, 2012, O'Hagan and Sasamoto, 2016), have also been used for quite some time and the sharing of results through subsequently published research has led to increasingly sophisticated experiments over the years. Research on fansubbing and fandubbing, started well over ten years ago (Díaz Cintas and Muñoz Sánchez, 2006), has had the merit of bringing to the fore special modes of AVT consumption, while also sparkling reflections on the increasingly active role of consumers (now better named prod*users*). Focusing mainly on perception, several scholars have been infusing psycholinguistic research into AVT, mainly focusing on accessibility (Fryer and Freeman, 2014) but certainly providing stimuli for many other scholars and areas of research.

As can be inferred above, all of these contributions to the study of AVT audiences have implied the development of ***multi***-, or ***inter***-disciplinary approaches to AVT (Di Giovanni, Agost and Orero, 2012), which can be said to have led audiovisual translation and media accessibility studies to maturity. Interdisciplinarity is steeped into audience research, whatever angle one chooses to adopt in a study. Below, a brief survey of experimental research on subtitle reception using eye tracking technologies (as in the study here reported) is offered.

2. **Eye tracking studies and subtitle reception**

If we shift from analysing audience research from a thematic perspective to discussing it in relation to the technologies used for this type of research, a pivotal role has long been played by eye tracking.

The use of eye tracking technologies in audiovisual translation research can nowadays be considered well-established and mature, having found applications in subtitling (from D'Ydewalle and Gielen, 1992, to Orrego Carmona, 2016, Kruger *et al.*, 2016, Doherty, 2018, and many more); audio description (Krejtz *et al.*, 2012, Di Giovanni, 2014); dubbing (Di Giovanni and Romero Fresco, 2019), and many other areas of investigation. First applied in relation to subtitling and

reading speed in the early Nineties, by psychologists such as Géry D'Ydewalle, eye tracking steadily entered into AVT research a few years after the turn of the new century, thanks to media accessibility scholar Pilar Orero and her international team working on the Digital Television for All project (https://cordis.europa.eu/project/rcn/191846/factsheet/es).

Focusing on subtitles for the deaf and hard of hearing (SDH) and audio description (AD), the project team used Tobii eye tracking technologies to evaluate reception across Europe. One of the outcomes of the Digital Television for All project was a book edited by Pablo Romero-Fresco,[2] where reception tests carried out in several European countries are reported, focusing on three forms of subtitles: edited, standard and verbatim. One of the aspects tested in this cross-national study was reading speed for subtitles, which has turned out to be indirectly central to the experiment presented below. In the book, Romero-Fresco reflects on the very definition of reading speed, pointing out that when it comes to subtitling, both intra- and interlingual, the expression may not be the most appropriate:

> The notion of reading speed could be used to refer to the reception of the subtitles by the viewers, but even then, it would not be thorough enough to account for the complexity of the audiovisual medium. Whereas in print the reading speed is set by the reader, who focuses mainly on words, in subtitling the speed is determined by the subtitler, while the viewer has to process both the subtitles and the images on the screen. […] print reading and subtitling viewing are completely different tasks. (2015: 337)

Romero-Fresco thus sets forth the notion of *viewing speed*, as a more accurate one to account for the process of viewing and understanding subtitles on film, and this notion proves particularly useful for our study reported below. Before focusing on eye tracking, a brief historical reflection on previous research on reading or viewing speed may here be of help, although it was not at the core of the initial hypothesis for the experiment here presented. Indeed, as we shall see, viewing speeds always have an impact on reception and they have to be a priority in the training of subtitlers.

[2] *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*, Bern: Peter Lang, 2015.

2.1 The early days of research on subtitle reading speed: from the 1970s to the turn of the century

As Romero-Fresco (2018) recalls, research on SDH (or captions) pioneered in the 1970s in the USA, mainly through a series of PhD dissertations which appeared some 10 years before subtitles were first aired on television in the same country. In 1973, as Romero reports, one large-scale experiment aiming to test reading speed rates was carried out:

> In 1973, Shroyer set out to obtain average reading rates from 185 deaf and hearing students in order to determine appropriate speeds for subtitling guidelines. Shroyer concluded that subtitles presented at 160 words per minute (wpm, the average speed for spontaneous conversation in English as found by Kelly and Steer in 1949) would exclude 84% of students with hearing loss in his sample, which led him to recommend 120wpm as an optimum speed for children. (2018: 201)

Research on reading speed rates, especially in relation to subtitles for the deaf, has been regularly carried out in the USA since then, making this country the leader in this field, with contributions from several disciplinary perspectives. In Europe, D'Ydewalle and a number of other scholars carried out systematic research on the same issue from the late Eighties, being indeed pioneers in the European landscape.

In 1996 and 1997, two, large-scale experiments were published by Carl Jensema, once again in the United States. Commissioned by the Described and Captioned Media Program (DCMP) funded by the U.S. Department of Education, these studies involved high numbers of participants (deaf, hard of hearing, hearing people) and different types of audiovisual texts. For his second study, a series of 24 short video segments were subtitled at different speeds (96, 110, 126, 140, 156, 170, 186 and 200 words per minute) and shown to 578 people. At the time of the second study, as Jensema himself recalls, there were already "over 500 hours of closed-captioned television programming shown each week" (1997: 2) in the U.S., and figures were on the increase. The experiment, which aimed at measuring "how comfortable people were with different caption speeds" (Ibid.), asked participants to rate on a 5-point scale if captions were *too fast* / *fast* / *ok* / *slow* / *too slow*. On average, the "Ok" speed was found to be associated with a value ranging from 140 to 156 words per minute, with interesting variations

across a spectrum of deaf, hard of hearing and hearing viewers. Here is one valuable remark by Jensema, useful for further studies: "Of particular interest was the adaptability exhibited by the respondents. As caption speed increased, the respondents recognized this, but most seemed able to adjust and did not appear to consider the captions unacceptable" (1997: 10). This seems to justify the steady increase of number of words per minute in subtitling which has been recorded internationally since those days. As we shall see in the following section, words per minute have come to be more commonly measured as characters per second, especially with reference to interlingual subtitling.

2.2 Subtitling, reading speeds and eye tracking: recent studies

As Romero-Fresco points out (2015: 336), viewing subtitles on film, understanding both the subtitles and the film, is not simply an act of reading: it is a complex cognitive act. As a matter of fact, more recent studies aiming to evaluate the reception of subtitles have focused on the notion of cognitive load (Gerber-Morón, Szarkowska and Woll, 2018) also breaking it down into three, different indicators: difficulty, effort and frustration. With a clear difference from te studies reported and/or carried out by Jensema and Romero-Fresco, who focused on SDH, Gerber-Morón, *et al.* centred their experiment on the syntactic segmentation of subtitles. They they found that cognitive load decreases systematically as subtitles are increasingly made 'readable' through apposite syntactic segmentation. Particularly interesting for the study presented below is the evaluation of cognitive load (effort) and comprehension by means of eye tracking experiments and questionnaires. In our experiment, interviews were not carried out, but several spontaneous declarations provided by the participants at the end of their test were recorded and proved extremely useful.

As a final reference to be highlighted in this section, another article by Szarkowska and Gerber-Morón (2018) is worth mentioning, as it focused on subtitle reading speed and relied on eye tracking. The authors aimed to evaluate comprehension, cognitive load, scene and subtitle recognition in films in two languages (Hungarian and English) and with speakers/readers of English, Polish and Spanish, once again focusing on interlingual subtitling. What is particularly interesting in this study, besides the different languages under investigation, is that the highest values of characters per second considered (12, 16 and 20 cps) are in line with the choices made by some of the most prominent

providers of streaming services such as Netflix for their interlingual subtitles. Although it is not easy to compare words per minute with characters per second, we could say that 12 cps would reflect the average value found by Jensema in his second study, i.e. 140-150 words per minute. As Szarkowska and Gerber-Morón point out, subtitle speeds have been on the rise for a long time, for both SDH and interlingual subtitling, although research on the latter is still more limited.

Their study relied on clips with interlingual subtitles at different speeds, created for the purpose of the experiment. Among the many interesting results obtained, the authors found that self-reported cognitive load, for the three parameters considered, yielded higher values for difficulty and effort but not significantly for frustration,[3] and that participants generally declared the lowest cognitive load for the slower (12 cps) subtitles across the three languages. Moreover, the authors observe that the English participants reported the greatest difficulties, which may be connected to their overall limited exposure to subtitles, given that AVT products are mainly produced in English with no need for interlingual subtitling. This finding is to be kept in mind when considering the results of our own experiment with Italian subtitles, as Italian viewers are similarly -if not equally- hardly ever exposed to subtitles.

3. **Testing viewing experiences across screens: *New Girl* with Italian subtitles**

In discussing their results, Szarkowska and Gerber-Morón observe that viewers in traditionally dubbing countries may generally find it more difficult to read subtitles, especially at high reading speeds, although a generally-held belief holds it that contemporary, young generations find it easier and more natural to read subtitles. Both issues were considered when setting up our experiment, whose main aim was to evaluate reception of both dialogue and images of the same audiovisual texts with Italian subtitles, experienced on screens of different sizes. A small team at the University of Macerata, Italy[4],

[3] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0199331#sec001 (accessed 10 December 2019).

[4] The team included Dr. Francesca Raffi and two Masters' students: Paolo Di Tosto

decided to test (20 to 30 years old) viewers' comprehension and overall reception of professionally-created subtitles for the popular Netflix series *New Girl*. This series was selected for several reasons: it is very popular with that age group and the various seasons have been on Netflix for a number of years. Also, *New Girl* is a comedy, and this features humorous exhanges and gags, accompanied by canned laughter. The protagonists are all young and the actions reflect their lifestyle. Two sequences from episode 1 of the third series of *New Girl* were selected, making sure that both contained whole scenes. This is the reason for a difference in overall duration of the clips: 1.54 minutes for the first and 2.17 minutes for the second.

Participants were selected to ensure a balance between men and women; language and translation students were excluded and participation of non-students was encouraged, so as to ensure a variety of competences, backgrounds and interests.

A movable, screen-based eye tracker, the Tobii X2, 60 Hz was used. It was mounted directly on a 24-inch TV screen and a 15-inch laptop screen, whereas for experiments using a 5-inch smartphone (Huawei P8) it was placed on a Tobii Mobile Device Stand 2, standard procedure for this type of experiment, with a Logitech C920 HD Pro webcam.

There were 20 participants for each screen group, for an overall of 60 participants. Taking into account the principles for spatial accuracy and precision discussed by Dalrymple *et al.* (2018), and aiming for full comparability of the three groups, we decided to discard 2 tests per group, to have this final setup:

- 18 participants for the 24-inch TV screen,
- 18 participants for the 15-inch laptop,
- 18 participants for the 5-inch smartphone.

All were in the selected age range, with an equal share of female and male participants for each group. They were asked to sit on a chair in front of the selected screen, and watch the videos as naturally as possible. The experiment lasted approximately 15 minutes for each participant, which included calibration time, the viewing of two clips

and Giada Ceruolo. The project was supported by SubTi Ltd, a U.K.-based audiovisual translation company interested in analysing the effectiveness of new and increasingly common viewing habits, such as watching audiovisual material on laptop and mobile phones.

and answering the two sets of questions which followed each clip viewing. Using Tobii Pro Lab software, the two clips were preceded by a black screen with a white text providing a brief context for the sequence to be watched.

3.1 The questionnaires: methods

As stated above, the main aim of this experiment was to assess viewers' comprehension of both verbal and visual elements in an American TV series subtitled into Italian and viewed on three differently-sized screens. Comprehension was evaluated through eye tracking tests and specific comprehension questions, to which general questions aiming to better frame overall reception were added. The attention was, therefore, on both verbal and visual intake, the verbal component relayed through interlingual, Italian subtitles. Viewing speed, as defined by Pablo Romero-Fresco, was indirectly central to the experiment. The almost verbatim subtitles provided by Netflix accompanied both clips virtually without a pause, and one of our hypotheses was that viewers would occasionally be unable to process both lines, as was later confirmed by eye tracking and questionnaire data. We had no clear-cut hypothesis on the incidence of viewing speed and long subtitles on reception through the different screens.

With the aim to obtain true-to-life results, a decision was made not to create ad-hoc Italian subtitles but use those offered by Netflix for *New Girl*. Selecting clips with dialogue of average density, with subtitles set by Netflix at a reading speed of 17 characters per second, with a limit of 42 characters per line, we ended up with long, almost verbatim subtitles. As we decided not to make any changes to the subtitles, the clips also presented instances of poor segmentation, incoherent use of punctuation and a few translation errors/ambiguities. The questionnaire began with basic demographic questions, asking participants to state their age, gender, occupation and to evaluate their knowledge of English, rating it on a 5-point scale as 1) poor, 2) not so poor, 3) fair, 4) good, 5) excellent. Then, the questionnaire comprised 8 questions for each clip, according to the following pattern:

- 4 questions aiming to evaluate intake/comprehension of visual elements,
- 4 questions aiming to evaluate intake/comprehension of verbal elements.

The questions were either closed (4 multiple-choice questions with 4 options each), or open (one of these four included an initial word to start from, with space for an open response). To wrap up, two questions, on a 5-point scale, asked participants to self-evaluate their comprehension of the dialogues and the images from 5 (excellent comprehension) to 1 (very limited comprehension).

3.2 The questionnaires: results

Before focusing on some of the data provided by the eye tracking tests and the major issues highlighted by participants, either as replies to the questionnaire or spontaneously at the end of the experiment, we will briefly focus on descriptive statistics for the two final questions.

	Please, evaluate your overall comprehension of the images (1 to 5)	Please, evaluate your overall comprehension of the dialogues (1 to 5)
TV	2.61	3.55
Laptop	2.43	3.00
Phone	2.31	3.16

Table 1: self-evaluations for comprehension of images and dialogues.

As can be seen in Table 1, the more 'traditional' screen for viewing audiovisual products has scored the highest values for self-reported comprehension of both images and text. This is a particularly meaningful finding, especially if related to the age range of the participants. Of great interest are also the figures obtained for the smartphone: increasingly familiar though it may be for today's viewers as a primary resource to enjoy audiovisual texts, it seems to have left our participants unsatisfied with their intake of visual information. We should recall here that these questions came at the very end of the experiment, i.e. after viewing both clips and replying to all comprehension questions. Therefore, replies are the direct result of the awareness elicited by the experiment itself. It is important to highlight that the questionnaire explicitly used the word "dialoghi" [dialogues] in the self-evaluative question. We decided not to specifically mention the subtitles, so as to leave the field open for

evaluation of the interaction between the original English speech and the Italian subtitles, which was in fact the object of direct comments by many of the participants.

In terms of overall correct or wrong answers to the 8 + 8 comprehension questions, Table 2 below provides a summary.

	Percentage of correct replies to comprehension questions	Percentage of wrong replies to comprehension questions
TV	39.5 %	60.5 %
Laptop	38 %	62 %
Phone	35.5%	64.5%

Table 2: Overall percentage of correct and wrong replies for the three screens.

These figures seem to corroborate those obtained from the self-evaluations reported above. The TV screen scored the highest percentage of correct replies, while the smartphone led to the most negative scores, for both image and dialogue comprehension.

3.3 Eye tracking data

For the sake of brevity, and to leave space for a discussion of key issues and areas of improvement in the training of subtitlers, this section aims at highlighting only some of the most valuable data obtained from the eye tracking tests which, as anticipated, were gathered for balanced groups of 18 individuals (9 male and 9 female) for each of the three screens.

In order to systematically compare data obtained for all three screens, several metrics were extracted from the overall experiment, namely: fixation duration (total and average), fixation count (total and average), visit count (total and average) and time to first fixation.

With regard to 'first fixation', this indicates the amount of time spent by viewers before fixating selected areas of interests (AOIs); and Table 3 below summarizes the data obtained for all three screens, in terms of time elapsed between the first visit on that screen and the first fixation on 1) the characters' faces (*faces*), and 2) the subtitles (*subs*). Values are expressed in seconds.

	Time to fist fixation on the AOI *faces* (average)	Time to first fixation on the AOI *subs* (average)
TV	0.26	7.2
Laptop	0.18	3.4
Phone	0.86	2.21

Table 3: Times to first fixations (in seconds).

The figures provided above, supported by the observation of each participant's gazeplot,[5] seem to highlight very plausible fixation patterns for the three screens. Starting from the TV screen, participants almost immediately fixated the characters' faces, thanks also to the fact that the faces were located almost at the centre of the screen. Subsequently, the viewers moved onto the subtitles after 7.2 seconds on average, which proves that they took time to scan and fixate the images before looking for meaning in the verbal written track. The figures are lower for the laptop screen, but even in this case participants fixated the character's faces almost immediately, to then move onto the subtitles after 3.4 seconds. Therefore, laptop viewing almost halves the time between the average first fixation on the faces and the subtitles as compared to viewing on the TV screen. As for the mobile phone, almost one second passed before participants fixated the characters' faces (+230% compared to the TV screen), which makes us think that some time was spent visually wandering outside the phone screen, which is certainly worth investigating further. Then, 1.35 seconds later, participants moved onto the subtitle block, with a decrease in the time spent on average on the images which amounts to 80.5% compared to the TV screen and 58% compared to the laptop screen.

To follow up on fixation patterns, one further metric worth analysing in detail here is the *fixation count*, related in the experiment to three AOIs, referring to the characters' faces (AOI *faces*), the overall subtitles (AOI *subs*), and the second line of the two-line subtitles (AOI *line_two*). It should be noted that the AOI *faces* is, on average, higher than the AOI for the subtitles, as 95% of the time both clips features

[5] A definition of gazeplot can be found here: https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/interpret/working-with-heat-maps-and-gaze-plots/ (accessed 10 December 2019).

more than one character on screen; and when only two characters are shown it is generally in close up. This point will have to be considered in more detail, to more accurately evaluate fixation count and duration in relation to AOI size.

As required for all experiments on moving images, the three AOIs for this study were captured dynamically, i.e. the AOI position on the screen changed shot by shot, following precisely the specific visual elements we were tracking for the duration of each clip.

As for *line_two*, a decision was made to create a specific AOI for the second line of each subtitle in order to see if, and to what extent, viewing speed, as set by Netflix for *New Girl*, truly allows viewers the time to read the second line. Table 4 below shows figures for the AOIs *faces, subtitles* and *line_two* with reference to the overall number of fixations, for both clips together. The average AOI per participant is given in brackets:

	Total fixation count *faces*	Total fixation count *subtitles*	Total fixation count *line_two*
TV	2086 *(115.8 p.p.)*	1418 (78.7 p.p.)	413 (22.9 p.p.)
Laptop	1925 *(106.9 p.p.)*	1487 (82.6 p.p.)	376 (20.8 p.p.)
Phone	1905 *(105.5 p.p.)*	1958(108.7 p.p.)	1389 (77.1 p.p.)

Table 4: Total fixation count (per group and per person).

As defined by a number of scholars, and also in the user's manual for Tobii Pro Lab,

> fixations are those times when our eyes essentially stop scanning about the scene, holding the central foveal vision in place so that the visual system can take in detailed information about what is being looked at.[6]

Fixations are recorded by eye trackers according to recording rate (Hz). They are clusters of gaze points, calculated by the eye tracker's own algorithm. Again, as stated in the Tobii Pro Lab manual, "Fixations are constructions, outputs of a mathematical algorithm that translates the sequence of raw gaze points into an associated sequence of fixations" (Ibid). Fixation length can vary, and can be visualized by

[6] https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/types-of-eye-movements/ (accessed 10 December 2019).

means of gazeplots or scanpaths. The duration of fixations can be influenced by several factors, including amount of visual intake, processing difficulties, etc. In relation to the number of fixations on faces found in our experiment, there is a progressive decrease across the three screens. This is more significant from the TV to the laptop than from the latter to the smartphone. On the other hand, the number of fixations on the subtitle block increases from the TV to the smartphone screen. In fact, on average, each participant fixated more on the smartphone subtitles and less on the laptop and TV screens. However, if we consider the overall fixation time for each AOI across the three screens (*total fixation duration*), which scores much lower values for the smartphone, the counts above suggest that fixations on the smartphone AOIs were generally shorter.

To further comment on the fixation counts, it may be useful to mention the number of subtitles in each clip, and their text distribution on either one or two lines.

- Clip 1: 39 subtitles, 29 on two lines and 10 on one line.
- Clip 2: 45 subtitles, 26 on two lines and 19 on one line.

In total, there were 84 two-line subtitles, and 55 one-line subtitles. Figures for subtitle fixations are quite revealing: for the TV screen, the average per person (p.p.) shows that approximately 7% of the subtitles were never fixated, whereas this percentage drops to 1.3% for the laptop screen. Viewers of the clips on the smartphone, on the other hand, scored on average +29% of fixations on subtitle blocks. Once again, these figures should be measured against average fixation durations to give us a better idea of the actual time spent on each AOI. Here, though, let us compare fixation times with the questionnaire replies. Overall, the highest score for correct understanding, both for visual and verbal information, is with the TV screen and lowest for the smartphone, which shows that a higher number of fixations does not necessarily correspond to better comprehension.

Scores for fixations on the second line of the two-line subtitles are also meaningful. More than half of the second lines were never fixated by TV screen viewers, a percentage that increases to approximately 62% for the laptop viewers. Smartphone users, on the other hand, fixated more, but again their overall comprehension was lower than the other two groups. Clearly, the screen size allows for much more saccades (eye movements), although this does not imply greater comprehension. As reported elsewhere (Di Giovanni, Romero Fresco,

2019), more fixations and saccades can in fact bear witness to a search for meaning.

4. Difficulty, effort, frustration

Although our experiment did not involve any direct measure of parameters such as difficulty, effort and frustration, which were analysed for instance by Szarkowska and Gerber-Morón, we will nonetheless consider these issues in relation to some of the results obtained, also reflecting on some of the spontaneous comments provided by participants upon completion of the experiment.

The limited eye-tracking data discussed above reveal that many participants, regardless of the screen format, did not manage to even look at the second line of subtitles and scored high levels of error in the comprehension questionnaire. As a matter of fact, 9 out of 18 (50%) participants for the TV screen group provided more incorrect than correct replies, with figures going up to 12 out of 18 (66.6%) for the laptop group and 14 out of 18 (77.7%) for the smartphone group. These results seem to suggest that viewing speeds set by Netflix for Italian viewers of *New Girl* are not to be considered fully appropriate - and inappropriateness often generates frustration.

Interestingly, 3 to 5 participants for each group provided spontaneous comments at the end of the experiment. Most of these comments revolved around the great difficulty of reading subtitles, which were said to be "too long", or "too difficult", often leading participants to attempt to retrieve meaning from the original audio track, thus causing even greater confusion and frustration. To our surprise, 6 participants went back to their English competence after completing the questionnaire, and told us they had probably overrated it at the beginning of the experiment. The words "difficulty", "tiring" and "attempts" were used by a number of participants in their spontaneous accounts, whereas frustration, although never spelled out, can be clearly inferred from such accounts.

As for the eye tracking data, besides those presented above, a qualitative analysis of gazeplots for each participant in the groups was performed. The gazeplot analysis shows that, in the presence of translation errors or ambiguous translation choices, individuals tend to produce more saccades (eye movements), especially in the form of regressions, backward movements over specific words or expressions. This seems to suggest that errors and ambiguities entail more cognitive

effort and also more time spent on the same, micro-verbal elements, with a consequent loss of pace in fruition. These data, paired up with the comprehension results, seem to reinforce the idea that awkward syntax and segmentation cause more saccades, regressions, and overall loss of comprehension (as proven by Gerber-Morón, Szarkowska and Woll, 2018). Falling outside the scope of this paper, further data analysis will be carried out elsewhere.

5. Training subtitlers today: inputs from reception studies

In the training of subtitlers, at least in my 15 years of experience in Italy and a few other countries, courses generally follow a top-down approach to the teaching of technical and linguistic parameters. Particular subtitling parameters are introduced, such as: number of lines, number of characters per line, reading speeds, use of punctuation and specific conventions (hyphens, italics, suspension marks, etc.), syntactic segmentation, shot and scene cuts. Examples are generally provided and students are encouraged to practice using videos of increasing difficulty, under the trainer's supervision.

However, the results of this experiment suggest that a bottom-up approach should be introduced and may well prove effective as an introduction to the subtitling course. For instance, exposing students to selected subtitled clips with different viewing speeds and asking them to report on their intake of visual and verbal information, both spontaneously and through comprehension questions, may elicit much more awareness about the importance of appropriate parameters in subtitling. Moreover, such an initial approach could be complemented with the presentation of reception experiments conducted with non-specialist viewers and using authentic subtitled materials, such as the one briefly analysed above, may prove useful in making students aware of many important aspects normally just discussed in class, like appropriate condensation in subtitle translation, for the sake of *viewing*, rather than simply reading subtitles on videos. As a matter of fact, one of the issues that subtitlers' trainers never tire of stressing is the importance of creating subtitles that accompany the viewing experience without becoming the main part of this experience. On the other hand, equally tantamount is the constant focus on translation quality, which many a reception experiment[7] has reported to be

[7] See, for instance, Di Giovanni: 2016.

essential for specialists as well as for the general audience. Translation errors or ambiguities, the use of inappropriate register in translation and the misuse of punctuation cause confusion and can generate frustration and dissatisfaction amongst the audience.

Language and translation competence must always remain central in the teaching of subtitling. Once again, a bottom-up approach is here useful, by exposing students to instances of poor or inappropriate translation directly, and by commenting on the results of reception experiments where difficulties and frustration are reported. Language and translation competence bring with them the importance of appropriate cultural adaptation in subtitling, which is certainly one of the most difficult elements to teach and learn. As Henri Béhar puts it, subtitling "is a form of cultural ventriloquism" (2004: 85), but the focus must remain, as smoothly as possible, on the puppet and not on the puppeteer. Teaching and learning how to dose cultural adaptation in subtitling is perhaps one of the most daunting tasks, one that can be targeted both through exposure to different instances of more to less appropriate cultural adaptation across audiovisual texts and genres, and also by presenting empirical research on the issue. Indeed, even though viewing speeds are very often imposed by commissioners of subtitles, such as in the form of pre-timed templates, teaching students to set appropriate times for subtitles, as well as the complex art of condensation, remains absolutely central.

To conclude, this article, and the experiment discussed, aims to support the crucial importance and potential of reception research for the training of subtitlers. It is suggested here that discussing experiments and their results, using experiment materials directly with the students (such as questionnaires) will boost awareness of the importance of all subtitling parameters and, above all, of translation quality.

References

Béhar, H. 2004. "Cultural ventriloquism". In A. Egoyan and I. Balfour (Eds.), *Subtitles. On the foreignness of film,* Cambridge, Massachussetts: Alphabet City Media/The MIT Press, 79-88.

Caffrey, C. 2012. "Using an eye–tracking tool to measure the effects of experimental subtitling procedures on viewer perception of subtitled AV content". In E. Perego (Ed.), *Eye Tracking in Audiovisual Translation*, Roma: Aracne, 223-258.

Di Giovanni, E., Agost, R., Orero, P. 2012. "Multidisciplinarity in audiovisual translation". In R. Agost, P. Orero and E. Di Giovanni (Eds.), *MonTi*, 4, 9-38.

Di Giovanni, E. 2014. "Visual and narrative priorities of the blind and non-blind: eye tracking and audio description". In *Perspectives. Studies in Translatology*, 22 (1), 136-153.

Di Giovanni, E. 2016. "Reception studies in audiovisual translation research: the case of subtitling at film festivals". In *Trans-kom*, 9 (1), 58–78.

Di Giovanni, E. 2018. "Dubbing, perception and reception". In E. Di Giovanni and Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation*, Amsterdam: John Benjamins, 159-178.

Di Giovanni, E., Romero Fresco, P. 2019. "Are we all together across languages? An eye tracking study of original and dubbed films". In I. Ranzato and S. Zanotti (Eds.), *Reassessing Dubbing. Historical Approaches and Current Trends*, Amsterdam: John Benjamins, 125-144.

Diaz Cintas, J., Muñoz Sánchez, P., 2006. "Fansubs: audiovisual translation in an amateur environment". In *JosTrans*, 6, 37-52, https://www.jostrans.org/issue06/art_diaz_munoz.pdf.

D'Ydewalle, G, Gielen, I. 1992. "Attention allocation with overlapping sound, image, and text." In K. Rayner (Ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading*, New York: Springer-Verlag, 415–427.

Doherty, S. 2018. "Analysing variable relationships and time-course data in eye tracking studies of translation processes and products". In F. Federici and C. Walker (Eds.), *Eye Tracking and Multidisciplinary Studies on Translation*, Amsterdam: John Benjamins, 71-94.

Fryer, L., Freeman, J. 2014. "Can you feel what I'm saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment". In A. Felnhofer, O.D. Kothgassner (Eds.), *Challenging Presence: Proceedings of the 15th International Conference on Presence*, Wien: Facultas.wuv. 99-107.

Gambier, Y. 2009. "Challenges in research on audiovisual translation". In A. Pym and A. Perekrestenko (Eds.), *Translation Research Projects 2,* Spain: Intercultural Studies Group, 17-25.

Gerber-Morón, O., Szarkowska, A., Woll, B. 2018. "The impact of segmentation on subtitle reading". In *Journal of Eye Movement Research*, 11 (4), DOI: 10.16910/jemr.11.4.2.

Jesema, C., McCann, R., Ramsey, S. 1996. "Closed-captioned television presentation speed and vocabulary." In *American Annals of the Deaf*, 141 (4), 284-292.

Jensema, C. 1997. *Viewer Reaction to Different Captioned Television Speeds".* Institute for Disabilities Research and Training, Weathon, MD, https://www.dcmp.org/caai/nadh30.pdf.

Krejtz, I., Duchowski, A., Szarkowska, A., Walczak, A. 2012. "Multimodal learning with audio description: an eye tracking study of children's gaze during a visual recognition task." In *Proceedings of the Association Computing Machinery (ACM) Symposium on Applied Perception*, ACM: New York, 83-90.

Kruger, J. L., Soto-Sanfiel, M., Doherty, S., Ibrahim, R. 2016. "Towards a cognitive audiovisual translatology: subtitles and embodied cognition". In R. Muñoz (Ed.), *Reembedding Translation Process Research*, Amsterdam: John Benjamins, 171-193.

O'Hagan, M., Sasamoto, R. 2016. "Crazy Japanese subtitles? Shedding light on the impact of impact captions with a focus on research methodology". In S. Hansen-Schirra and S. Grucza (Eds.), *Eyetracking and Applied Linguistics*, Berlin: Language Science Press, 31–58, doi:10.17169/langsci.b108.234.

Orrego Carmona, D. 2016. "A reception study on non-professional subtitling. Do audiences notice any difference?". In *Across Languages and Cultures*, (17) 2, 163–181.

Pavesi, M. 2013. "*This* and *that* in the language of film dubbing: a corpus-based analysis. In *Meta*, 58 (1), 107-137.

Pavesi, M. 2014. "The Pavia corpus of film dialogue: a means to several ends". In M. Pavesi, M. Formentelli and E. Ghia (Eds.), *The Languages of Dubbing. Mainstream Audiovisual Translation in Italy*, Bern: Peter Lang, 29-55.

Pavesi, M. 2019. "Corpus-based audiovisual translation studies: Ample room for development". In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation Studies*, London/New York: Routledge.

Romero Fresco, P. 2018. "Reception studies in live and pre-recorded subtitles for the deaf and hard of hearing". In E. Di Giovanni and Y. Gambier (Eds.), *Reception Studies and Audiovisual Translation*, Amsterdam: John Benjamins, 199-224.

Romero Fresco, P. (Ed.). 2015. *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern: Peter Lang.

Szarkowska, A., Gerber-Morón, O., 2018. "Viewers can keep up with fast subtitles. Evidence from eye movements". In *PLoS ONE*, 13 (6), https://doi.org/10.1371/journal.pone.0199331