

Aleksandar Kavgić*
Faculty of Philosophy
University of Novi Sad

UDC 81'25: 791
DOI: 10.19090/gff.2019.2.17-38
Originalni naučni rad

COMPREHENSION OF CONTENT AND RETENTION OF VISUAL DETAILS: A NEW METHODOLOGY FOR SUBTITLING AND DUBBING**

The canonical view of subtitling and dubbing assumes that they differ in relation to the viewer's experience of cinematic illusion. There is however very little empirical evidence that subtitled videos are detrimental in this respect. This pilot empirical research was devised to create and test the methodology which can fill this empirical gap in audio-visual translation (AVT). The empirical research involved 68 volunteers divided into two groups watching 3-minute video clips. One group watched a video dubbed in Serbian, another subtitled in Serbian with Arabic audio, and, finally, a poorly subtitled video in Serbian with Arabic audio. The other group watched the same videos, but they were subtitled in Serbian with Arabic audio, dubbed in Serbian, and accurately subtitled in Serbian with Arabic audio, respectively. The participants were then quizzed to assess content comprehension of the text featured in dubs and/or subtitles, and for retention of visual details. The results seem to indicate that the methodology provides valuable insights into the problem, that the canonical views on differences between subtitling and dubbing are inaccurate, and that good subtitling has a positive effect on both content comprehension and on attention to visual details, thus improving the overall viewing experience.

Key words: audio-visual translation, content comprehension, visual details, subtitling, dubbing, English, Serbian, empirical research

1. INTRODUCTION

Audio-visual translation (AVT) is a sub-field within the field of Translation Studies (TS) and deals with translation of text in multimedia, which is which is extremely important for the consumption of media productions in an increasingly

* kavgić@ff.uns.ac.rs

** This paper is part of the research on Project No. 178002, entitled *Languages and Cultures in Time and Space*, which is funded by the Serbian Ministry of Education, Science and Technological Development.

globalized community (e.g. Díaz-Cintas, 2010). Within AVT itself, subtitling (Cintas, 2013) and dubbing (Chaume, 2013) are the most widely used and accessible modes of AVT, unlike audio description, voiceover and surtitling, which are rarer (Baños–Piñero, & Cintas, 2015). The importance of AVT for the consumption of multimedia by the viewers stems from its influence on how a multimedia product – be it an episode in a series, video blog or even a video commercial – is perceived and received by its target audience, i.e. viewers (Gambier, 2009: 52–56): when a visual media production is well dubbed or subtitled, it is more accessible to the audience and provides the desired effect, and when a piece of visual media production is poorly dubbed or titled, this may confuse or disappoint the audience, perhaps to the point of turning the production off. In that sense, AVT is considered to be a necessary, but suboptimal distraction from the multimedia production, in the sense that it reduces cinematic illusion and distracts from the images (Koolstra–Peeters, & Spinhof, 2002). However, “very few studies have dealt with the issue of reception in AVT, and even fewer have looked at empirically” (Gambier, 2009: 52)

This paper aims to fill the aforementioned gap in our understanding of the reception of AVT and to do it by means of empirical research designed to assess if subtitled videos indeed are detrimental to the visual experience, if poorly done subtitles have an even more detrimental effect, and if it is indeed true that dubbed videos seem to promote the visual experience. More specifically, 68 students from the Department of English Studies, Faculty of Philosophy, University of Novi Sad, Serbia, were subject to an empirical study based on the combination of a control group research design and an incomplete counterbalanced measures research design: they watched a combination of three dubbed and subtitled videos and were quizzed after each video with two sets of questions which assessed content comprehension of the text featured in dubs and/or subtitles and the retention of visual details featured in the clips, respectively. Thus, the aims of the research were to determine the following: 1. the differences between subtitling and dubbing in the comprehension of text i.e. linguistic content, 2. the differences between subtitling and dubbing with respect to images and/or cinematic illusion, and 3. the assumption that poor subtitling necessarily affects these same qualities.

The research presented in this paper should be understood primarily as a pilot study, whose aim is to attest the validity of the methodology devised for the chosen topic. In that respect, the paper proceeds by stating the theoretical considerations and a detailed outline of the methodology, while the empirical study

results and concluding remarks, each in a separate section, provide the validation for the devised methodology and present the case for using it in further studies.

2. THEORETICAL CONSIDERATIONS

In line with the goals, as indicated in the introduction, the theoretical framework of the research is squarely within the domain of descriptive translation studies (DTS) which represent “a descriptive, empirical, interdisciplinary, target-oriented approach to the study of translation” that “has branched out in several directions including technical translation, audiovisual translation or interpreting, among others” (Rosa, 2010: 94). More specifically, DTS has the goal “to explore its [translation’s] context and its conditioning factors, to search for grounds that can explain why there is what there is” (Hermans, 1999: 5). In that sense, DTS is inherently compatible with the empirical approach to real-world research – as chosen in this study – where the empirical approach is understood to have the aim “to manipulate a treatment [or independent] variable (i.e. an intervention) to determine the effect on a dependent variable (outcome)” (Leatherdale, 2019: 21).

However, it should be emphasized that one of the main theoretical considerations in the research is that of subtitling norms. Namely, as will be described in the methodological section, one of the independent variables was the size, timing, duration and positioning of subtitles, which were purposefully made non-compliant with subtitling norms in order to empirically examine if they have any influence on comprehension of either visual or linguistic content. For the purpose of empirically and purposefully deviating from the norms, it was necessary to determine what the subtitling norms are. Despite a considerable diversity in subtitling practices around the world (e.g. Díaz-Cintas & Remael, 2007), a very convenient summary of the norms can be found in Karamitroglou (1998), which also contains a set of recommended practices for subtitling. In short, the subtitling norms (Karamitroglou, 1998) imply that no subtitle should have more than two lines, where each line should have no more than 35 characters: the subtitle should be placed in the lower part of the screen with the first line shorter than the second one (in order to minimize the negative impact on the visual content) and should not remain on the screen for less than 1 second (a single line) or more than 7 seconds (two lines). Additionally, the syntactic segmentation of subtitles should be done at the highest nodes and there should be no hyphenation at the end of the line.

The most important theoretical consideration for the research is that subtitles are detrimental to the transfer of information, impede information processing because they reduce the visible area of the original picture, distract

attention away from the picture, take more mental effort than watching a dubbed program, educe cinematic illusion and distract from the images (Koolstra et al., 2002: 327–334). Moreover, subtitles are said to disturb the unity of picture and image, but also look unnatural as they are asynchronous (i.e. the subtitled words do not appear one by one as they are uttered, but multiple words appear at once) (Koolstra et al., 2002: 344).

3. RESEARCH QUESTIONS AND HYPOTHESES

In full alignment with the theoretical considerations outlined above, the research was designed to answer one general, overarching question, comprised of three specific questions. Most explicitly, the main research question was to determine the real effects of choosing subtitling over dubbing, i.e. how this choice may contribute to or detract from (which constitute two of the three specific questions), a) comprehension of linguistic information contained in the video, and b) respect of images and cinematic illusion. The third specific question was to determine if poor subtitling (from the normative point of view) has the same effects as quality subtitles, or it further augments the negative effect of subtitling. In short, as indicated in the introduction, the research aims to determine if watching a subtitled video indeed is inferior in terms of comprehension and experience to watching a dubbed video.

Having in mind the research questions, the initial hypothesis was that, despite a lack of studies dealing with the issue of AVT reception, the referenced literature is correct; i.e. subtitling is detrimental both to the comprehension of linguistic content and to cinematic illusion, and poor subtitling will augment the detrimental effects of subtitling in general.

4. RESEARCH METHODOLOGY

The research presented in this paper was designed as an empirical one, resembling an experiment with a single variable in each segment. More precisely, the research employed a combination of the standard control-group design and incomplete counterbalanced measures design, where the research participants were two groups of university students: 68 subjects altogether.

At this point, it is essential, for the sake of complete transparency, to clarify that “counterbalanced measure design”, as used in this paper, refers to an empirical research where “as with the standard repeated measures design, the researchers want to test every subject for both conditions”, but, instead of having a traditional

control group and experimental group, “they divide the subjects into two groups and one group is treated with condition A, followed by condition B, and the other is tested with condition B followed by condition A” (Shuttleworth, 2009). The advantage of the counterbalanced measure design is that it allows for having more data even with smaller sample sizes. Visually, a basic counterbalanced measure design with two conditions is shown in Figure 1 below.

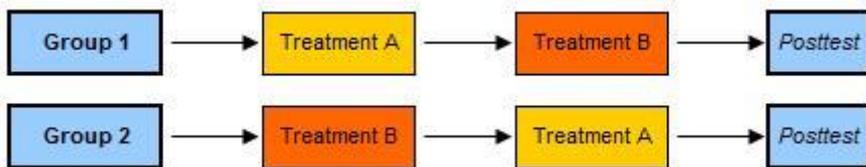


Figure 1: An example of a counterbalanced measures design when there are two possible conditions, A and B (taken from Shuttleworth, 2009)

The research was designed so that each of the two groups watched 3 video clips of approximately 3 minutes each: one video with voiceover, one which was properly subtitled, and one which was poorly subtitled. After each video, the subjects were asked to answer a set of 10 questions consisting of two smaller subsets: (1) questions 1–5 about visual details, mostly colors, with varying levels of difficulty (these questions were about objects that were either centrally or peripherally placed with the time on screen between 2 and 4 seconds, depending on the level of difficulty), and (2) questions 6–10 about non-visual content, i.e. things that had to be processed cognitively. The questions and their types are shown in Table 2, Table 3 and Table 4. Participants answered these questions via a quiz functionality in Moodle. The videos were also hosted in Moodle.

Both the video playback and quiz were automatized (automatic playback with no pauses) and timed (quizzes), so that all participants had exactly the same amount of time to watch the videos and answer the questions about them. It should be noted that the subjects participated in the research by physically coming to the research’s office and sitting in front of a PC with access to Moodle. In other words, the research design purposefully did not allow for any possibility of remote access. This decision was made so as to ensure the researcher’s supervision of the empirical part of the research and to prevent any deviations from the envisaged procedure, although it

made the research more difficult organizationally and prolonged the duration of data collection. A sample of the instructions to the participants and a quiz question is given in the screenshots presented in Figures 2 and 3. The Serbian text in Figure 2 explains that the quiz is timed and that there is between 12 and 15 seconds between each two questions before the quiz automatically moves forward. Finally, it should be noted that for questions about colors, only black and white photos/stills were used, as exemplified in Figure 3.

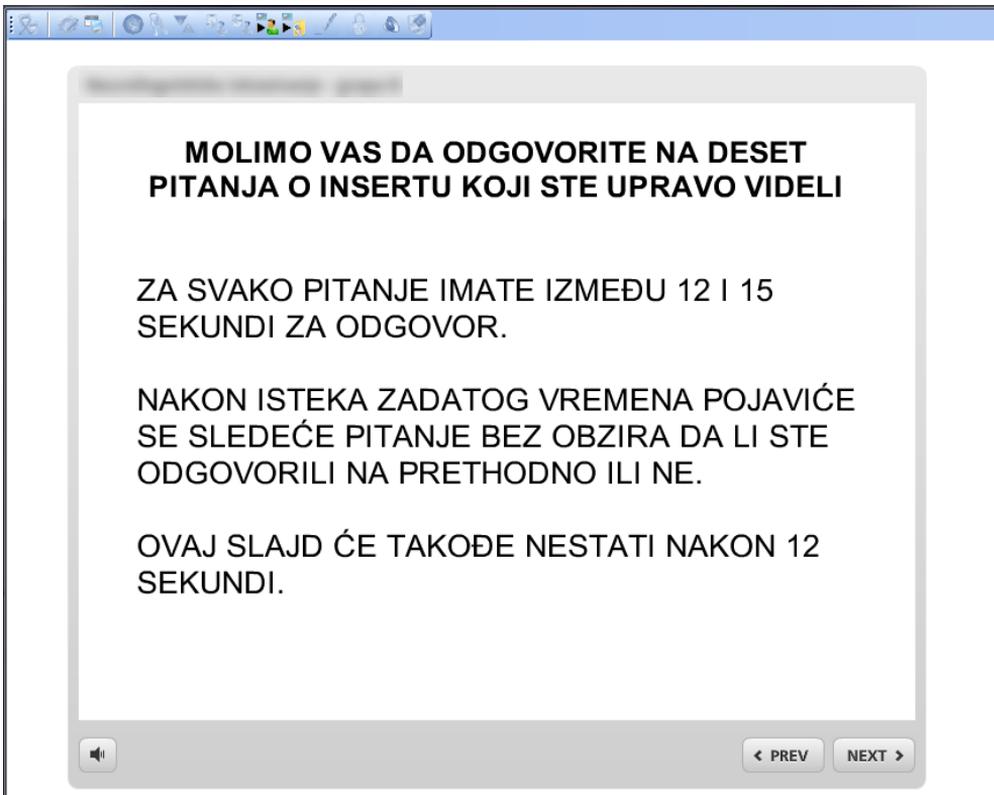


Figure 2: A screen clipping of instructions displayed to the research participants



Figure 3: A screen clipping of a quiz question.

The videos which were chosen for the research were extracted from two BBC documentaries: *Blue Planet* and *Life of Mammals*. The clip from *Blue Planet* was about the blue whale, while the clips from *Life of Mammals* were about the oceanic adaptations of mammals and about the impact of humans on the environment.

4.1. RESEARCH VIDEOS

There were two sets of videos used in the research, with one set consisting of two pairs of videos and the other of a single pair of videos. The sets had different functions: the first set was used to assess the influence of the two main modes of AVT – subtitling and dubbing – on the maintenance of cinematic illusion and comprehension of content, while the second one was employed to assess the influence of the normative validity of subtitles on the comprehension of content.

The first set of videos (henceforth: Set 1) consisted of two subsets: (1) a 3-minute clip about the blue whale from the *Blue Planet* series, with a voiceover in Serbian (and no subtitles), for group A, and a voiceover in Arabic and good subtitles in Serbian, for group B, and (2) a 3-minute clip about the oceanic adaptations of mammals from the *Life of Mammals* series, with good subtitles in Serbian and a voiceover in Arabic), for group A, and a voiceover in Serbian and no subtitles, for group B. Set 1 can be considered a counterbalanced measure design, which was the approach which was taken in order to maximize the sample size, as the primary focus of the research was the investigation of the effects of subtitling vs. dubbing on the comprehension of the presented content. Specifically, it is Set 1 which was designed to assess the influence of the subtitling vs. dubbing choice on the maintenance of cinematic illusion and comprehension of content.

The second video set (henceforth: Set 2) may be said to constitute a traditional experiment, where group A is the control group (native speakers of Serbian watching a video clip with voiceover in Serbian), while group B is the experimental group (native speakers of Serbian watching a video clip with Serbian subtitles but with voiceover in Arabic). For Set 2, the decision was made not to take a counterbalanced measure design, as a bigger sample size was not necessary and would make the procedure longer. Most explicitly, Set 2 was designed to investigate the secondary question of the research: the influence of the normative validity of subtitles on the comprehension of content. Using a counterbalanced approach for this would make the empirical procedure over 20 minutes long, which was not in line with the participants' request not to spend more than 15 minutes. Had the research had the possibility to provide the monetary incentive to the participants, Set 2 would have been designed as a counterbalanced measure design – however, given that the participants were volunteers, the research had to accommodate for their request of a maximum duration of about 15 minutes.

At this point it is important to emphasize two methodological decisions that have been made, and provide clarification, and/or justification, for them: one relates to the use of Arabic for voiceover, and the other has to do with the use of different videos in the two subsets.

The voiceover in Arabic was chosen for all subtitled videos in the research, for a simple reason that most students taking part in it were students of English whose English language competences were at least at the C1 level, according to the Common European Framework of Language Reference. Prior to accepting to take part in the research, the students were asked if they spoke any Arabic and those who did were not accepted. The reasoning for this was that, given the purpose of Set 1 –

to assess if the voiceover does indeed promote cinematic illusion – the only way to control this variable was to make sure that the audio track could not be used for comprehension by the participants first or secondary language, but that only subtitles could be used for comprehensions.

Set number:	Group A	Group B	Purpose
Set 1	Blue Planet – The Blue Whale (Serbian voice over, no subtitles) Life of Mammals – Ocean (Arabian voice over, good Serbian subtitles)	Blue Planet – The Blue Whale (Arabic voice, good Serbian subtitles) Life of Mammals – Ocean (Serbian voice over, no subtitles)	assess the influence of subtitling vs. dubbing choice on the maintenance of cinematic illusion and comprehension of content
Set 2	Life of Mammals – Homo Sapiens (bad Serbian subtitles)	Life of Mammals – Homo Sapiens (good Serbian subtitles)	assess the influence of the normative validity of subtitles on the comprehension of content.

Table 1: The organization of research videos in two sets: contents and purpose

The use of two different videos in Set 1 was not the optimal one but was necessary in order to prevent the priming of participants. Priming is understood to refer to the phenomenon in which exposure to one stimulus influences responses to subsequent stimuli of the same kind (with no conscious guidance or intention). More specifically, had the same video been used first with Arabic voiceover (with Serbian subtitles) and then with Serbian voiceover (with no subtitles) in Group A, and then in Group B with Serbian voiceover (with no subtitles) and then with Arabic voiceover (with Serbian subtitles), it is very likely that after the first video was shown the participants would be primed by the quiz to focus on details of the video that may have eluded their attention in the initial watching. For this reason, the decision was made to use different videos, but with exactly the same length and amount of words. In short, the organization of videos was made in a manner shown in Table 1.

Figures 4 and 5 below show examples of poorly- and well-made Serbian subtitles in Set 2. It is important to mention that the blue progress bar was not visible to the participants.



Figure 4: Examples of poorly-made Serbian subtitles in Set 2



Figure 5: Examples of well-made Serbian subtitles in Set 2 (the blue progress bar was not visible to the participants)

4.2. QUIZ QUESTIONS AFTER EACH VIDEO

The quiz questions, designed to provide insights into the comprehension of the contents and the level of cinematic illusion, were separated into two groups: 1) questions that assess the comprehension of non-visual (i.e. linguistic) content, and 2) questions that assess focus on visual details (i.e. maintenance of cinematic illusion). It should be noted that the term comprehension is understood here in its dictionary definition of “the ability to understand something”. The original questions were in Serbian because the participants were all native speakers of Serbian, and the focus of the research was the use of localized content in AVT. The questions are presented in Tables 2, 3 and 4 below.

Question	Serbian question	Translation of the Serbian question into English	Type of question
S1q1a	Da li se na početku inserta, pre pojavljivanja naziva serije, u nekoj od scena vide delfini?	Can dolphins be seen at the beginning of the video clip, before the series title is displayed?	Yes/no
S1q2a	Koje boje je kanta koja se vidi na sredini inserta?	What color is the bucket which appears towards the middle of the video clip?	Open
S1q3a	Označite klikom deo palete koji nije boje drveta.	Click on the part of the floating palette which does not have the natural color of wood.	Open
S1q4a	Koje boje je odbačena ribarska mreža?	What is the color of the discarded fishing net?	Open
S1q5a	Koje boje su mehurovi koji održavaju algu na površini?	What is the color of the floats that keep the algae afloat?	Open
S1q6a	Ribe sa početka inserta zovu se ?	The fish from the beginning of the clip are called?	Multiple choice
S1q7a	Ovaj plutajući predmet je?	This floating object is?	Multiple choice
S1q8a	Ove veće ribe koje ljubomorno čuvaju svoj plutajući stan zovu se?	The bigger fish that jealously guard their floating home are called?	Multiple choice
S1q9a	Ovaj ogromni busen morske trave otkinuo se sa priobalne stene u ?	The huge chunk of sea plant has been torn from a coastal rock in?	Multiple choice
S1q10a	Ovaj insert je po Vašem mišljenju iz epizode koja se bavi?	This video clip is from an episode whose main topic is?	Multiple choice

Table 2: The list of questions used in the quiz after the first research video in Set 1 (i.e. Set1a).

Question	Serbian question	Translation of the Serbian question into English	Type of question
S1q1b	Da li se tokom inserta u nekoj sceni vidi čitav kostur plavog kita?	Is the entire skeleton of the blue whale ever fully visible in the clip?	Yes/no
S1q2b	Koje boje je košulja Dejvida Atenboroa?	What color is David Attenborough's shirt?	Open
S1q3b	Koje boje je kitov jezik ?	What is the color of the whale's tongue?	Open
S1q4b	Koje boje je Dejvidov čamac?	What is the color of David's boat?	Open
S1q5b	Koje je boje Dejvidov pojas za spasavanje?	What color is David's life jacket?	Open
S1q6b	Plavi kit je dugačak?	How long is the blue whale?	Multiple choice
S1q7b	Kapacitet kitovih pluća je?	What is the blue whale's lung volume?	Multiple choice
S1q8b	Kitovo srce u minuti udara ?	The blue whale's heart beat rate is?	Multiple choice
S1q9b	Ove male kosti su ostatak ?	What are these small bones the remnants of?	Multiple choice
S1q10b	Ovaj insert je po Vašem mišljenju iz epizode koja se bavi?	This video clip is from an episode whose main topic is?	Multiple choice

Table 3: The list of questions used in the quiz after the second research video in Set 1 (i.e. Set1b).

Question	Serbian question	Translation of the Serbian question into English	Type of question
S2q1	Da li se tokom inserta u nekoj sceni vidi Spejs šatl u letu iznad Zemlje?	During the clip, can the Space Shuttle be seen flying above Earth.	Yes/no
S2q2	Koje boje je kišobran Dejvida Atenboroa?	What color is David Attenborough's umbrella?	Open
S2q3	Koje boje je trup aviona u ovoj sceni ?	What color is the plane's fuselage in this scene?	Open
S2q4	Koje je boje kombajn u ovoj sceni ?	What is the color of the harvester in this scene?	Open
S2q5	Koje boje su majice sakupljača čaja u ovoj sceni?	What is the color of tea harvesters' shirts in this scene?	Open
S2q6	Kolika površina zemlje se koristi za poljoprivrednu proizvodnju?	How much of Earth's landmass is used for agricultural production?	Multiple choice
S2q7	Gde se nalazi pustinjska plantaža sa početka inserta?	Where is the desert plantation from the beginning of the clip located?	Multiple choice
S2q8	Fundamentalni razlog za napredak ljudske tehnologije bio je?	The fundamental reason for the advancement of human technology was?	Multiple choice
S2q9	Prema Dejvidu Atenborou, ako ne promenimo odnos prema prirodi, može nas snaći sudbina koje drevne civilizacije?	According to David Attenborough, if we do not change our relationship with nature, we may endure the same fate as what ancient civilization?	Multiple choice
S2q10	Ovaj insert je po Vašem mišljenju iz epizode koja se bavi?	This video clip is from an episode whose main topic is?	Multiple choice

Table 4: The list of questions used in the quiz after the third research video in Set 2.

As it can be seen from Tables 2, 3 and 4, virtually all questions that assess the viewer's focus on visual details were based on color. There were three reasons why this decision was made. Firstly, color is widely regarded as the essential visual stimulus and experience (Adams, & Osgood, 1973). Secondly, color is the key cognitive information channel which plays a major role in memory and enhances memory performance (Wichmann–Sharpe, & Gegenfurtner, 2002). Thirdly, and most importantly, recent research studies suggest that color is the key element for visual memory (a nice review and summary of multiple studies can be found in Dzulkifli & Mustafar, 2013), which makes it the best choice for assessing the impact of subtitling and dubbing on retention of cinematic illusion. In other words, if a particular mode of AVT has a negative influence on color as the core element of visual memory, it may be confidently assumed that that mode of AVT has a negative influence on the cinematic illusion as a whole.

4.3. EMPIRICAL PROCEDURE

The empirical procedure consisted of welcoming the participant and ushering them into the office. The participants were then asked to sit in front of a 27" LCD monitor on which an interactive Moodle page with videos was open in

full screen. The participants were asked not to adjust the brightness or any settings of the monitor, the volume of the speakers, which was set to 100% (2W stereo speakers), or to move them. The interactive Moodle page then guided the participants through the empirical procedure, while the researcher was sitting silently in the corner of the office, making sure that the conditions were identical for all participants. The results were stored in Moodle and then exported as CSV for further statistical processing.

5. RESEARCH RESULTS

The physical limitations of the paper in terms of its length, as well as its pilot nature, necessitate that the presentation of the research results be general and similar to that from a bird's eye perspective. The word 'general' should here be interpreted in the sense of not going into detailed discussions about each question, but focusing on the big picture of having, or not having, statistically significant results.

The general results of the empirical research are presented in Table 5, while Table 6 summarizes the statistical findings for t-test and Table 7 for Pearson Chi-square test. To be more precise, despite a relatively small sample size (68 subjects: 31 subjects in group A and 37 in group B), the methods of descriptive statistics were applied to the dataset. In particular, the Pearson chi-square test was used, and it was cross tabulated across groups. The choice of the statistical test was dictated by the type of data generated by the research. More explicitly, the questions were either open ended or multiple-choice and generated categorical data, meaning that the t-test (which requires two variables: a categorical one with exactly two levels, and a quantitative one which is estimable by a mean) could not be used (Freedman–Pisani, & Purves, 2007: 488–494). Having this in mind, the Pearson chi-square, which operates on categorical variables that may have any number of levels, had to be chosen (ibid: 523–539). The spatial limitations of the article do not permit the presentation of the variety of answers provided for each question, but, for example, the answers for S1q5a (the color of floats) included several varieties of “žute“ (yellow, adjective) “žut“ (yellow, adjective) and “žuta“ (yellow, noun), but also various inaccurate answers such as “plave“ (blue, adjective), “zelene“ (green, adjective), “narandžaste“ (orange, adjective), “crvene“ (red, adjective) and “ne znam“ (I don't know): this also illustrates the need to use the Pearson chi-square test.

Prior to the Pearson chi-square (Table 7), a t-test was done. The t-test was performed separately for the means of visual- and content-related questions in each

set and for each set as a whole. The results turned out to be statistically not significant at either $p < .10$, $p < .05$ or $p < 0.01$ confidence intervals. The t-test results are shown in Table 6a and Table 6b, which show group statistics and independent samples t-test results, respectively.

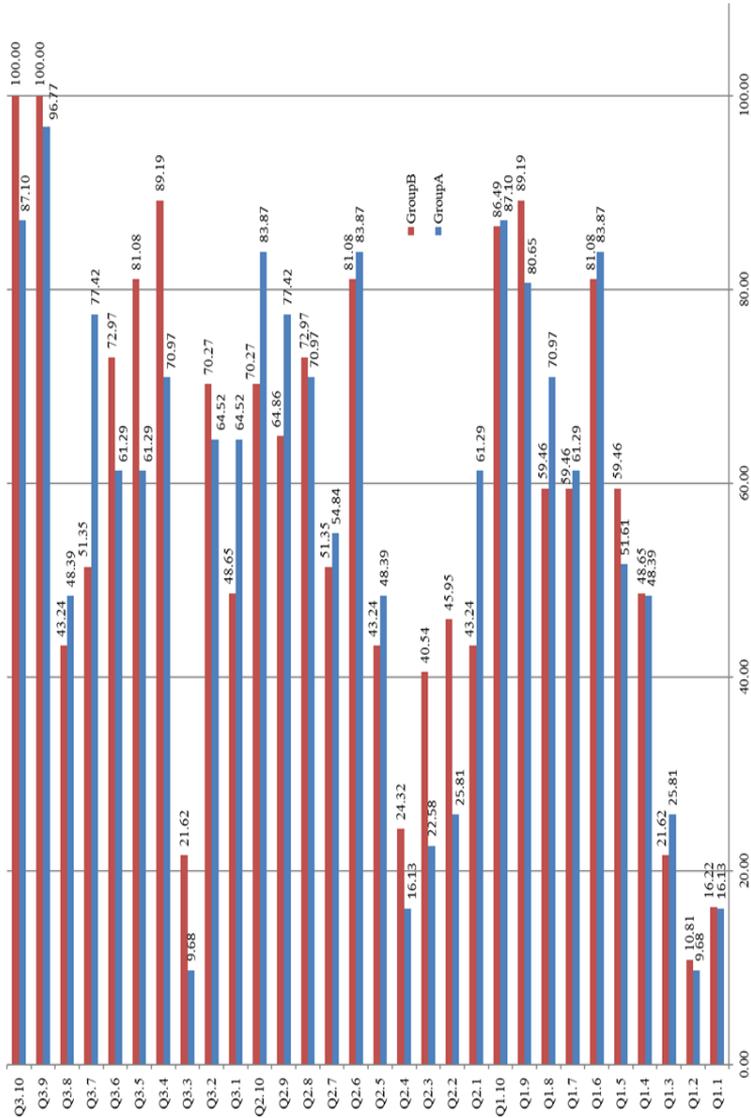


Table 5: The general results of the research: results for all video pairs showing the percentage of correct answers in each group: Q1.1-Q1.10 are quiz questions for the first video pair in Set 1, Q2.1-Q2.10 for the second video pair in Set 1, while Q3.1-Q3.10 are quiz questions for the only video pair in Set 2.

	Group:	Number of questions	Mean (correct answers)	Std. Deviation	Std. Error Mean
Visual Questions					
Set 1	Group A	5	30.1980	18.93079	8.46611
	Group B	5	31.3520	21.41812	9.57847
Set 2	Group A	5	34.8400	19.13843	8.55896
	Group B	5	39.4580	8.67587	3.87997
Set 3	Group A	5	54.1960	25.13233	11.23952
	Group B	5	62.1620	27.29639	12.20732
All 3 sets of visual questions together	Group A	15	54.1960	25.13233	11.23952
	Group B	15	62.1620	27.29639	12.20732
Content Questions					
Set 1	Group A	5	76.7760	10.55277	4.71934
	Group B	5	75.1360	14.60508	6.53159
Set 2	Group A	5	74.1940	12.06885	5.39735
	Group B	5	68.1120	11.03360	4.93438
Set 3	Group A	5	74.1940	19.48722	8.71495
	Group B	5	73.5120	26.50976	11.85553
All 3 sets of content questions together	Group A	15	75.0547	13.54701	3.49782
	Group B	15	72.2533	17.49794	4.51795
Visual and Content Questions (all sets)					
All 3 sets of visual and content questions together	Group A	30	57.3997	25.57316	4.66900
	Group B	30	58.2887	24.77761	4.52375

Table 6a: Group statistics based on the % of correct answers in Group A and B in all three sets of questions.

		Levene's Test for Equality of Variances				t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
		Visual Questions									
									Lower	Upper	
Set 1	Equal variances assumed	.296	.601	-.090	8	.930	-1.15400	12.78367	-30.63319	28.32519	
	Equal variances not assumed			-.090	7.881	.930	-1.15400	12.78367	-30.71078	28.40278	
Set 2	Equal variances assumed	6.531	.034	-.491	8	.636	-4.61800	9.39734	-26.28831	17.05231	
	Equal variances not assumed			-.491	5.577	.642	-4.61800	9.39734	-25.04149	18.80549	
Set 3	Equal variances assumed	.184	.679	-.480	8	.644	-7.96600	16.59353	-46.23076	30.29876	
	Equal variances not assumed			-.480	7.946	.644	-7.96600	16.59353	-46.27604	30.34404	
All sets	Equal variances assumed	.400	.532	-.547	28	.589	-4.57933	8.37014	-21.72480	12.56613	
	Equal variances not assumed			-.547	27.950	.589	-4.57933	8.37014	-21.72618	12.56751	
		Content Questions									
Set 1	Equal variances assumed	2.164	.179	.204	8	.844	1.64000	8.05815	-16.94214	20.22214	
	Equal variances not assumed			.204	7.282	.844	1.64000	8.05815	-17.26693	20.54593	
Set 2	Equal variances assumed	.063	.808	.832	8	.430	6.08200	7.31297	-10.78173	22.94573	
	Equal variances not assumed			.832	7.937	.430	6.08200	7.31297	-10.80525	22.96925	
Set 3	Equal variances assumed	.735	.416	-.046	8	.964	.68200	14.71407	-33.24871	34.61271	
	Equal variances not assumed			-.046	7.346	.964	.68200	14.71407	-33.78184	35.14584	
All sets	Equal variances assumed	.976	.332	.490	28	.628	2.80133	5.71372	-8.90270	14.50536	
	Equal variances not assumed			.490	26.347	.628	2.80133	5.71372	-8.93587	14.53853	
		Visual and Content Questions									
All sets	Equal variances assumed	.016	.901	-.137	58	.892	-.88900	6.50107	-13.90230	12.12430	
	Equal variances not assumed			-.137	57.942	.892	-.88900	6.50107	-13.90258	12.12458	

Table 7b: Independent samples t-test based on the % of correct answers in Group A and B in all three sets of questions.

The statistical analysis for Pearson Chi-square was conducted “just in case” after the negative t-test results, but the researcher did not assume that the analysis would yield any statistically significant results. However, this assumption proved to be wrong, as there were 5 statistically significant results in the entire set of 30 questions: one statistically relevant result in the first set of questions (Table 2) and four statistically significant results in the third set of questions (Table 4). The summarized results for the Pearson chi square test cross-tabulated across groups is given in Table 8.

Pearson Chi Square values (cross-tabulated across groups):	Question:
0.083	S1q2a (color of the bucket)
0.055	S2q4 (color of the harvester)
0.069	S2q5 (color of tea pickers' shirts)
0.026	S2q7 (the location of the plantation)
0.024 f	S2q10 (the topic of the last video)

Table 8: The list of questions for which answers were statistically significantly different between the groups.

The result for S1q2a is quite significant because it is the first question in the entire empirical research that assesses whether the AVT mode influences cinematic illusion and attention to visual details. Specifically, although the rest of the questions regarding attention to visual details do not show statistically significant differences, they do show (see Table 5) conspicuous differences between the groups. Namely, visible, but statistically not significant, differences between groups in subsequent questions regarding attention to visual details may be the consequence of priming. Plainly speaking, the subsequent answers may be a consequence of a participant thinking hard about the color of different objects from the video. However, the most important conclusion that can be drawn from S1q2a is that the higher score was obtained from Group B, i.e. the group who listened to the Arabic voiceover and read Serbian subtitles: not in Group A who listened to the Serbian voiceover, as was the initial hypothesis and as the canonized view assumes. Generally, Group B scored higher on most questions in this empirical research (see Table 5). In other words, the pilot research results indicate that the additional cognitive load necessary to process subtitles may have, as a kind of side effect, heightened attention to visual details and overall better comprehension of the video. S1q2a is the only statistically significant result in Set 1, i.e. among the questions designed to test ‘visual comprehension’.

On the other hand, the remaining four statistically significant results all come from Set 2, which was designed to test the influence of the normative quality of the subtitles on visual and textual comprehension. The S2q4 and S2q5 results, where the participants were asked about the color of the harvester and the color of tea picker's shirts, and where Group B (good subtitles) scored significantly better than Group A (poor subtitles), indicates that large subtitles which appear longer than they should may indeed be detrimental to visual details. In short, poor subtitles may hinder the processing of visual details if they occupy too much space on the screen. The normatively valid and invalid subtitles for this scene (S2q4) are shown in Figure 6 below.



Figure 6: The scene targeted by the question S2q4: invalid subtitles (Group A) are in the left screenshot and valid ones (Group B) in the right screenshot.

The results of S2q7 and S2q10 may initially seem puzzling, as Group A, with poor subtitles, scored better on the question about the location of the plantation (S2q7), while the overall comprehension, measured by S2q10, which had to do with the topic of the video, was better in Group B, i.e. with good subtitles. However, it should be noted that in S2q7 the poor subtitle targeted by the question actually implied that the word “Arizona” (the correct answer) appeared on the screen for 11 seconds (instead of the recommended duration of up to 6 seconds). In short, almost double the exposure to the word, paired with the awkwardness of looking at the subtitle while there is no voice in the audio track, probably contributed to Group A participants remembering the location of the plantation better. On the other hand, in S2q10 the general hypothesis for Set 2 is confirmed: poorly done subtitles that violate the good subtitling practice seem to have a negative effect on the overall comprehension of the non-visual content, as a statistically significant number of participants from Group A (poor subtitles) could not recognize the general topic of the video (contrary to the 100% score of Group B).

6. CONCLUDING REMARKS

This study was designed to test if a combination of the standard control-group design and incomplete counterbalanced measures design in the empirical approach can be used to fill the gap in assessing the differences between subtitling and dubbing in the comprehension of text i.e. linguistic content, with respect to images and/or cinematic illusion, and in determining if poor subtitling necessarily affects these same qualities. Although the t-test results for the overall results obtained by the two groups were not statistically significant, the Pearson chi-square analysis (cross-tabulated across the groups) identified 5 statistically significant results, which may be taken as evidence that, pursuant to further optimizations, the methodology devised for this research can be used for large-scale research in the differences between subtitling and dubbing as two main modes of AVT.

The five statistically significant results seem to indicate that some of the canonical views on the differences between subtitling and dubbing are not necessarily accurate. Specifically, it seems that the use of subtitles increases attention to visual details, most likely as the brain is involved in increased visual processing due to the presence of subtitles. On the other hand, the effects of poorly done and normatively invalid subtitles do fit with the initial hypothesis: poorly done subtitles are detrimental to cinematic illusion and overall comprehension of the topics covered by the video. Having that in mind, this pilot study also seems to

indicate that an increased duration of a subtitle's visibility on the screen can be used to improve memorization of facts mentioned in the video.

Needless to say, as a pilot study, this research has merely scratched the surface of the general topic of the influence of the AVT mode on textual and visual comprehension of the contents contained in the video. Therefore, the main contribution of the study is that it has validated the selected methodology. On the other hand, further studies are needed to truly assess the general topic of the paper and draw general conclusions on the two main modes of AVT: subtitling and dubbing.

Aleksandar Kavgić

RAZUMEVANJE SADRŽAJA I PAMĆENJE VIZUELNIH DETALJA:

NOVA METODOLOGIJA ZA TITLOVANJE I SINHRONIZACIJU

Rezime

U prevodilačkim studijama i oblasti audiovizuelnog prevođenja postoji uvrežen i kanonski stav da se titlovanje i sinhronizacija razlikuju u meri u kojoj gledalac može da uživa u kinematografskoj, tj. vizuelnoj iluziji, pri čemu se smatra da titlovanje negativno utiče na kinematografsku iluziju i vizuelni doživljaj, dok sinhronizacija na iste ima pozitivan uticaj. S druge strane, gotovo da ne postoje empirijski dokazi koji potkrepljuju ove tvrdnje. Upravo zbog toga, za ovaj rad se može reći da je proistekao iz nastojanja da se stvori i potvrdi empirijska metodologija koja može da proveri validnost ovih uvreženih tvrdnji i stvori osnove za dalja istraživanja u ovakvim tematskim okvirima. Ovo probno istraživanje zasnovano je na empirijskom istraživanju u kome je učestvovalo 68 ispitanika podeljenih u dve grupe koje su gledale kratke, trominutne inserte iz prirodnačkih dokumentarnih serija BBC-ja, pri čemu su nakon svakog inserta odgovarali na pitanja koja proveravaju pamćenje vizuelnih detalja i opšte razumevanje pogledanog sadržaja. Istraživanje je kombinovalo standardnu metodologiju sa eksperimentalnom i kontrolnom grupom, ali je primenjeno i kontrabalansirano istraživanje, što je učinjeno u cilju povećanja broja odgovara, tj. radi stvaranja većeg skupa podataka. Jedna grupa je gledala insert sinhronizovan na srpski bez titlova, insert sa titlovima na srpskom i sinhronizacijom na arapskom i insert sa normativno loše napravljenim titlovima na srpskom i sinhronizacijom na arapskom. Druga grupa gledala je iste inserte, ali je prvi bio sinhronizovan na arapski sa srpskim titlovima, drugi je bio sinhronizovan na srpski bez titlova, a treći je imao normativno kvalitetno napravljene titlove na srpskom i sinhronizaciju na arapski. Analiza odgovora ispitanikâ na pitanja o pamćenju vizuelnih detalja i razumevanju sadržaja otkrila je više statistički značajnih razlika među grupama, što predstavlja dokaz validnosti metodologije razvijene za istraživanje. Štaviše, statistički značajne razlike među grupama ukazuju da konvencionalni stavovi možda nisu

utemeljeni, jer se čini da titlovanje pozitivno utiče na pamćenje vizuelnih detalja, kao i na razumevanje sadržaja inserata.

Ključne reči: audiovizualno prevođenje, razumevanje sadržaja, vizuelni detalji, titlovanje, sinhronizacija, engleski, srpski, empirijsko istraživanje.

REFERENCES

- Adams, F. M. – Osgood, C. E. (1973). A Cross-Cultural Study of the Affective Meanings of Color. *Journal of Cross-Cultural Psychology*, 4(2), 135–156. <https://doi.org/10.1177/002202217300400201>
- Baños Piñero, R.– Cintas, J. D. (2015). *Audiovisual Translation in a Global Context Mapping an Ever-changing Landscape*. Retrieved from <http://public.eblib.com/choice/publicfullrecord.aspx?p=4082298>
- Chaume, F. (2013). Research Paths in Audiovisual Translation: The Case of Dubbing. In C. Millán & F. Bartrina (Eds.), *The Routledge handbook of translation studies* (pp. 306–320). Routledge.
- Cintas, J. D. (2013). Subtitling: Theory, Practice and Research. In C. Millán & F. Bartrina (Eds.), *The Routledge handbook of translation studies* (pp. 273–285). London ; New York: Routledge.
- Díaz-Cintas, J. (2010). Media for All: New Developments. In J. Díaz-Cintas, A. Matamala, & J. Neves (Eds.), *New Insights into Audiovisual Translation and Media Accessibility* (pp. 11–22).
- Díaz-Cintas, J. – Remael, A. (2007). *Audiovisual Translation: Subtitling*. Manchester, UK; Kinderhook, NY: St. Jerome Pub.
- Dzulkifli, M. A. – Mustafar, M. F. (2013). The Influence of Colour on Memory Performance: A Review. *The Malaysian Journal of Medical Sciences : MJMS*, 20(2), 3–9.
- Freedman, D. – Pisani, R. & Purves, R. (2007). *Statistics*, 4th Edition. New York: W. W. Norton & Company.
- Gambier, Y. (2009). Perception and Reception of Audiovisual Translation: Implications and Challenges. In H. C. Omar, H. Haroon, & A. Ghani (Eds.), *The Sustainability of the Translation Field* (pp. 40–57). Kuala Lumpur: Malaysian Translators Association.

- Hermans, T. (1999). *Translation in Systems. Descriptive and System-Oriented Approaches Explained* (Vol. 45). Manchester: St. Jerome Publishing.
- Karamitroglou, F. (1998). A Proposed Set of Subtitling Standards in Europe. *Translation Journal*, 2(2), 1–15.
- Koolstra, C. M. – Peeters, A. L. & Spinhof, H. (2002). The pros and cons of dubbing and subtitling. *European Journal of Communication*, 17(3), 325–354.
- Leatherdale, S. T. (2019). Natural experiment methodology for research: a review of how different methods can support real-world research. *International Journal of Social Research Methodology*, 22(1), 19–35. <https://doi.org/10.1080/13645579.2018.1488449>
- Rosa, A. A. (2010). Descriptive translation studies (DTS). In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of translation studies* (Vol. 1, pp. 94–104). Amsterdam: John Benjamins.
- Shuttleworth, M. (2009, May 8). Counterbalanced Measures Design - Counterbalancing Test Groups. Retrieved May 03, 2019, from Explorable website: <https://explorable.com/counterbalanced-measures-design>
- Wichmann, F. A. – Sharpe, L. T. & Gegenfurtner, K. R. (2002). The contributions of color to recognition memory for natural scenes. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 28(3), 509–520.