Exploring the Educational Potentials of Language Learning with Netflix Tool: An Eye-Tracking Study

Aikaterini Gouleti, Giorgos Dimitriadis, and Michalis Kokonis

Aristotle University of Thessaloniki, Greece.

Abstract

Digitization has revolutionized the industry of home entertainment services, enhancing the audience’s viewing experience and offering an abundance of choices within a new intercultural and multilingual reality. This pluralistic environment is a yet uncharted terrain of resources that could be exploited while popular culture and school curricula reach a decisive juncture. Evaluating the potentials of using new media learning tools, developed in line with the expansion of Over-the-top media services, has underpinned the objectives of this research. The analysis of eye movement data depicts the viewing patterns on three versions of the same film extract via Netflix streaming service. The first was screened with standard interlingual subtitling and the other two were viewed via Language Learning with Netflix (LLN) platform, a newly launched tool which allows the simultaneous, dual presentation of both the original dialogue and its translation. This paper aims to explore the proliferation of accessible options among different modes of audiovisual language transfer within an online participatory environment. In the emergent new media culture, the educational potentials of bilingual subtitling can challenge well-established borderlines and habit formations of viewership.

Keywords: new media, audiovisual language transfer, bilingual subtitling, streaming services, reception studies.

https://doi.org/10.26262/exna.v0i4.7920

Introduction

The constantly evolving world of home entertainment, especially with the rise of Video on Demand (VoD) services in the last decades, has already taken the world by storm offering new opportunities within novel types of viewing experience. Parallel to studying the audiovisual products themselves, attention has been extensively drawn to the way viewers interact with the screen and the multimodal nature of the material. Language, as an intrinsic part of films which goes hand in hand with image and sound, is turning into a focal point for multimedia and multilingual communication. The unprecedented proliferation of options among audiovisual translation methods and language pairs, provided by these services, sends interdisciplinary and reception-oriented studies into uncharted territory. The present experimental research is an attempt to delve deeper into the viewers’ minds by drawing empirical data literally from their eyes.

1 This research is carried out/funded in the context of the project “Implementing Audiovisual Media in Education: Evaluation and Application of Eye-tracking Data.” (MIS 5005088) under the call for proposals “Supporting Researchers with Emphasis on New Researchers” (EDULLL 34). The project is co-financed by Greece and the European Union (European Social Fund - ESF) by the Operational Programme “Human Resources Development, Education and Lifelong Learning 2014-2020.”
Using eye-tracking technology, gaze data were collected from young viewers (11-13 years) while watching three versions of the same film extract via Netflix streaming service. The first version was screened with standard interlingual subtitling and the other two were viewed via Language Learning with Netflix (LLN) platform, a newly launched tool that allows the simultaneous, dual presentation of both the original dialogue and its translation. The aim of this research is thus twofold: Firstly, we seek to draw findings on how the viewers/learners may refocus their attention in the presence of gaze attractors, such as written language content of different types. Secondly, the analysis of pre-and-post-viewing language tests contributes to the second objective of this research, which is to explore the educational potentials of Netflix in terms of language learning and apply both traditional and technological heuristic means to investigate both the challenges and the opportunities offered by the new media.

An Abundance of Choices

Up until very recently, countries had been characterized according to their preferred mode of audiovisual language transfer as subtitling, dubbing or voice over countries. Firmly impeded habit formation in the audiences worldwide had been dictating the perpetuation of the same preferences they originally opted for, for the years to come. This classification that was formed decades ago due to political or financial reasons (O’Sullivan and Cornu 21) is now revisited since, in the new era of globalized cultural settings, the once clear-cut boundaries are now blurred. The advent of DVD disclosed early of the new game-changing reality with the proliferation of accessible language choices between different subtitled or dubbed versions. The new digital video streaming platforms can now adjust to the viewers’ demands, allowing them to switch between subtitled audiovisual content in their own native language, the language of the original soundtrack or other. The audio track is also customizable, and viewers can now choose among a wide range of different options for watching the program of their preference in different dubbed versions. The national preferences still adopted by public TV broadcasters in different countries are now being challenged by new media. This abundance of choices is greatly appreciated by the better informed and foreign-language literate contemporary viewership that has been exposed (via studies, work or entertainment) to this new multilingual reality.

The latter is further enhanced in practice since viewers are presented with a wide range of international audiovisual content to choose from. Over-the-top video service platforms enrich their catalogs with Asian and European productions that are now creating an ever-increasing intercultural tank of diverse audiovisual products.² Apart from the need to fulfill the viewers’ demands for multicultural content, regulations have also been imposed to support this cultural diversity as in the case of the European audiovisual sector, where “Members of the European

² To name but a few, the German Dark (2017), Money Heist (2017) from Spain, and the Hindi Sacred Games (2018) have been included in the lists of most-viewed Netflix TV series worldwide. In the film industry, Parasites (2019) is the first non-anglophone language film to win the Oscar for Best Picture and this fact alone denotes the emergence of a new reality in the world of entertainment.
Parliament (MEPs) have passed legislation to ensure that before long, 30% of content in the catalogues of video-on-demand platforms’ should be European.\textsuperscript{3}

Under these circumstances, the demand for audiovisual translation skyrockets and occupies a central role in both defining and determining the new standards. The field now attracts the attention of scholars and practitioners who are seeking to explore the consumption patterns of audiovisual translated productions. An increased volume of content, a unique variety of genres, different language pairs and formats, and new software tools are unquestionably posing new demands and open up new perspectives. Subtitling, both as a mode of professional video localization and as a practice nowadays accessible even to amateurs worldwide, is further examined in relation to its impact on intercultural communication, including its educational perspective.

**Subtitling in Participatory Culture**

The exponential growth in video circulation\textsuperscript{4} has triggered participatory subtitling practices aimed at rendering foreign language video content accessible to audiences worldwide. Subtitling has been tightly linked to different forms of participatory culture. The latter have been presented by media theorist Henry Jenkins (“Confronting the Challenges of Participatory Culture”\textsuperscript{8}) and their multifaceted dimensions can be found in a diverse list of paradigms. Jenkins uses participation as a term “that cuts across educational practices, creative processes, community life, and democratic citizenship” and delineates the forms of participatory cultures that may include “affiliations, collaborative problem-solving, expressions and circulations”\textsuperscript{5}.

Beginning with the subtitling-related practices, participation is originally observed in the form of affiliations. Membership as an act of creating and sharing one’s creations can have a formal or institutional structure as in the case of Amara Subtitling Platforms,\textsuperscript{5} organized by the Participatory Culture Foundation.\textsuperscript{6} Subtitling is also a common practice among informal affiliations organized in the form of online fan-subbing groups,\textsuperscript{7} where Jenkins’s notion of “mentorship usually passed from experienced members to novices” is indeed applicable. In both


\textsuperscript{5} We are referring to the non-profit section of Amara (the one including volunteer subtitling for videos on TED Talks, GitHub, Udacity, Scientific American etc.), available at amara.org/en/subtitling-platform since from 2012 there has been an “Enterprise Solutions” section open to corporate clients (Baker 463).

\textsuperscript{6} A non-profit organization whose mission is to render internet-circulated information accessible worldwide is available at [pculture.org/](http://pculture.org/)

\textsuperscript{7} Different terms are used to describe the act of non-professional subtitling whereby amateur practitioners work individually or usually in teams to create and then upload subtitles for the programs of their preference. These volunteer translators’ activities have been referred to as “community translation,” “social translation,” “volunteer translation,” and “collaborative translation” among other labels that tend to overlap (Orrego-Carmona 1).
In this way, consumers themselves are now able to participate in the process of making and creating—in this case, translating—media content and contributing to the circulation of audiovisual material and experiences.

_Collaborative Problem-solving_ is also conducted through activities offered online through a social networking web platform through programs, such as Clipflair (Foreign Language Learning through Interactive Captioning and Revoicing of Clips), whereby teachers and learners can work together and share their ideas, complete tasks and develop new knowledge. Contrary to “passive media spectatorship” (Jenkins, “The Cultural Logic of Media Convergence” 93), people with similar interests across the globe can work together to eradicate language barriers. Their goals may significantly vary, and their _expressions_ can range from pure entertainment (e.g. subtitling the popular genre of Japanese Anime) and social activism,\(^9\) to education\(^10\) (e.g. subtitling academic-oriented videos for Massive Open Online Courses/MOOCs) and, more specifically, to foreign language learning. People that now see subtitling as a tool to enhance communication in the ever-expanding digital ecosystem, find themselves in need of circulating their contributions since they believe in their significance in this pluralist and linguistically accessible new media landscape. Even in the field of social media, specific tools, such as _YouTube_ Video Manager and _Facebook_ Captions Editor, are embedded to facilitate users in the subtitling process of uploaded videos which renders them linguistically accessible to worldwide viewership.

In the present paper, we focus on a relatively new type of subtitling, that of bilingual/dual/double subtitling (Toda 161). It must be clarified that bilingual subtitles have been produced in such geographical areas where two languages are spoken, as in Belgium, Finland, Jordan, and Israel (Diaz-Cintas and Remael 18); additionally, bilingual subtitles are common practice in international film festivals in order to attract wider audiences (19). There were also occasional examples of bilingual subtitles in supplementary video materials used as language learning resources. These cases excluded, the simultaneous appearance of subtitles in two languages has been a rare finding to this day.

For the purposes of this research, we will examine a new course taken in the practice of bilingual subtitling which emanates as a unique example of utilizing the affordances of new media and has emerged hand in hand with the advent of the emerging media landscape. From the incentives that lie behind its creation, to the circulation of the tools that facilitate its implementation, bilingual subtitling can be regarded as part of new media textualities. After an

\(^8\) The program is available at [clipflair.net](http://clipflair.net/).

\(^9\) Luis Perez-Gonzales talks of the transition of subtitling as a form of mediation in the media landscape from a representational to an interventionist practice (Audiovisual Translation 58).

\(^10\) The _Turn-on the Subtitles_ campaign is about enhancing reading and fighting illiteracy through the use of same-language subtitling. It has recently gained huge visibility throughout the world. You can find out more at [turnonthesubtitles.org](http://turnonthesubtitles.org/). Accessed 12 Mar. 2019.
Exploring the Educational Potentials of Language Learning with Netflix Tool: An Eye-Tracking Study

117

initial presentation of its potential language gains, this paper will focus on the eye-tracking study conducted to investigate how young viewers/learners responded to it in terms of their visual attention distribution and the parameters of foreign language learning.

**Bilingual Subtitling and Language Learning with Netflix**

Experimenting with the different ways in which bilingual or dual subtitles can be used by viewers to help them enhance their language learning is relatively new. Internet users have been exchanging ideas on how to use the technology at hand in order to achieve this end-result manually. During the last decade, we have been witnessing an exponential increase in the open-source programs or *Google Chrome* extensions that allow for the simultaneous presence of two sets of subtitles both in the original language of the video (intralingual)\(^{11}\) and in the viewers’ native language (interlingual). A possible explanation for the popularity of bilingual subtitles is the fact that they are thought to combine the benefits of intralingual and interlingual subtitles. The former display the written version of the spoken word, enhancing the acquisition of new vocabulary; the latter provide the translation of the spoken dialogue and contribute to the overall comprehension of the video (García 470).

To our knowledge, all online tools that involve bilingual subtitling are drawing on its pedagogical character and educational value and are promoting its integration in foreign language learning. Fleex player\(^{12}\) and DualSub\(^{13}\) were two of the forerunners. Viewers can use these open-source tools to merge two subtitle files creating a bilingual set of subtitles. Making use of other types of participatory culture exchange (such as fansubs), the subtitles are automatically collected from web sites where user-generated materials are being exchanged. If, however, viewers have just one input subtitle file, the program allows them to participate in the translation process, creating the target language version of their choice themselves, acting as not mere consumers of cultural products but (similar to amateur subtitlers) as producers-consumers or prosumers (Pérez-González 337), contributing thus in the making and remaking of cultural artifacts.

As it is observed, their interface can be customized in terms of subtitles’ positioning or appearance (Fig.1), adjusting to the users’ demands. Extra options, such as the interactive greying-out of words or popping up of vocabulary definitions while hovering over words, are also available (Fig.2). Acting as web-based viewing tools, these platforms “aim to maximize vocabulary learning while ensuring that the learner fully understands the video and enjoys watching it” (Kovacs and Miller 1).

---

\(^{11}\) Intralingual subtitles have been used primarily to cater for the needs of deaf and/or hard of hearing.

\(^{12}\) For more details visit fleex.tv/en/Home/Index4

\(^{13}\) For more details visit bonigarcia.github.io/dualsub/index.html
Some of these options are available to more than 300,000 users that have so far installed the Language Learning with Netflix (LLN) tool which, despite being in its infancy and compared to other platforms (officially launched in December 2018) at least, has received the greatest attention and media coverage. It runs as a Chrome extension on which audiovisual content streamed through Netflix can be screened with subtitles shown in two languages, allowing viewers to pair and compare translation in their own language with the original text and dialogue. It is “a sofa-based language lab,” according to The Guardian, providing a dynamic way of harnessing the educational potential of Netflix which broadcast its content in 26 languages in 190 countries; a “clever service” that makes use of Netflix’s massive catalog and all the major languages in which it already offers subtitles. According to the latest statistics (Fig.3), in the final quarter of 2019, Netflix had over 167 million paying streaming subscribers worldwide as well as over 4.6 million free trial customers. The following chart clearly depicts the exponential increase, since the number of subscribers has almost doubled in the last three years.

---


Developing the software was a joint venture undertaken by David Wilkinson and Ognjen Apic, who experimented with making their own language learning software before collaborating on Language Learning with Netflix. Apic launched his first venture with the “crude” predecessor of LLN, called Lingoplayer, and then joined forces with Wilkinson to take the next step. It is a clear case of how collaborations can be fostered between people who share the same interests and meet through online channels of participatory culture in order to create open-source technology accessible to audiences worldwide. As they state, they have worked independently and have not contacted Netflix until now, but have relied on users' circulations of their materials. In their effort to disseminate and circulate their project online, developers themselves have used social media and popular blogs where they initiated discussions with end-users to get feedback on their work and troubleshoot the shortcomings.

There are several options available for LLN users, as they can choose between displaying two types of translations: machine-created and human. As stated in their online guide, the former is very literal, “helping you understand the structure of the sentence in the language you are studying.” Nevertheless, errors can be frequently encountered using this option, whereas human translation is often useful, “as it correctly translates expressions and idioms.” Adjusting playback speed, video auto-pause and highlighting words to get their translation, pronunciation and examples on how to use them in context are some of the extra features offered.

After implementing online open-source tools, the users/members of participatory cultures are also expected to express their opinions on the degree of their satisfaction with them. There has been some negative criticism expressed through blogs or tweets emphasizing that this type of language learning methods cannot yet substitute accredited language classes.

Language Gains from Bilingual Subtitles

Subtitling is a challenging process as it requires the use of the audience’s cognitive resources to process information from different channels. Gottlieb notes the simultaneous interplay of multiple sources of information, specifically non-verbal elements (such as background music and/or sound effects and other audio) and moving image (visual) as well as verbal elements (such as the spoken words heard in the audio track) and the subtitles (visual) displayed on screen.

17 The same developers have also launched a beta version chrome extension for LLY (Language Learning with Youtube), a sister extension to LLN (December 2019). For more information visit chrome.google.com/webstore/detail/language-learning-with-yo/jkhhdaafjenpmmpcpgdjkhffdpmmcb.
18 Lingoplayer is an open source program creating bilingual subtitles (automatically translated by Google Translate). It is available at oaprograms.github.io/lingo-player/.
19 One post thread that includes such discussions can be found at reddit.com/r/languagelearning/comments/9ornvu/learning_languages_with_netflix_chrome_extension. Accessed 20 Jan. 2019.
20 Since they updated the extension to v3.0.0, machine translation and word saving are now paid features.
21 An indicative example where users express their doubts on the effectiveness of this type of language learning methods can be found at reelwords.ca/perhaps-word-nerding-through-netflix/. Accessed 19 Mar. 2019.
Apart from balancing out this concurring intake, further challenges need to be overcome since, due to the dynamic nature of the materials, the audience has no control over the presentation rate of these sources of information. Replacing the monolingual subtitles with bilingual ones can set additional cognitive demands and potentially influence the aspired learning outcomes since viewers need to process extra information. Bilingual subtitling combines reading and listening to the dialogue in the source language as well as reading the translated version in the viewers’ mother tongue.

The potential benefits of watching audiovisual material with intralingual or interlingual subtitles for foreign language learning have been investigated by many scholars so far. The effectiveness of subtitles is still a topic of much debate and subject to various conditions, such as the context of use, the subtitle type (interlingual, intralingual, reversed,22 or bilingual), and the relationship between the language of the soundtrack and that of the subtitles (Liao et al. 1). Especially as far as vocabulary acquisition is concerned, scholars seem to agree that watching both captioned (in the same-language) and interlingually subtitled video materials can help learners enhance different types of vocabulary acquisition, such as recall or recognition (Sydorenko; Vanderplank; Baltova; Bianchi and Ciabattoni). Similarly, Stephen A. Bird and John N. Williams note that bimodal presentation leads to enhanced vocabulary learning (501). Noa Talavan, also, refers to the “mnemonic power of images . . . enhanced by the presence of sound and text,” emphasizing how important visual associations are in memory for vocabulary learning (5). What is more, Susan Neuman and Patricia Koskinen suggest that “captioned television appears to provide a particularly rich language environment which enables students to learn new words incidentally through context” (104).

Taking all these widely researched benefits of monolingual subtitles into account, we further investigate the potential positive effects of bilingual subtitles on vocabulary learning. Liao et al.’s recent eye-tracking study dismisses the concern that the increased amount of text on screen in the condition of bilingual subtitling causes higher extraneous cognitive load (92). The authors have also found that the increased redundancy (between the subtitle text and the spoken dialogue, as well as between the subtitle text and the image) in bilingual subtitles does not have a negative impact on the learners’ performance. Researching the pedagogical implications of the use of differential subtitles, Mingyue Li states that bilingual subtitles have proved to be more effective than monolingual subtitles with regard to students’ vocabulary acquisition in both the short and the long term (143).

In order to shed more light on how the presence of bilingual subtitles tampers with the allocation of viewers’ visual attention and, consequently, with the learning gains, an experimental study was conducted based on a mixed-method approach. The physiological instrument of eye-tracking was used as an online measure in order to gain insights on how the

22 In the case of reversed subtitles, the audio track is in the viewers’ native language and the subtitles are translations in the foreign language.
learners’ attention is being distributed over the visual display of both subtitles and non-verbal visual input. Measuring the number and duration of eye fixations can deliver significant results since, as noted by Julie Sedivy, “the simplest assumption underlying the use of eye movement monitoring for psycholinguistic study is that the people tend to direct their eye gaze to things they are attending to in their visual environment” (116). What is important is that, with the advent of these new methods, we can now study how viewers process the information intake in real-time. Rayner notes that “longer fixations are widely argued to indicate either difficulty in extracting information or that the object is more engaging in some way” (155). Undeniably, many elements compete for the viewers’ attention and act as gaze attractors and, with the aid of eye-tracking measurements, the distribution of learners’ attention is monitored in the different subtitled video conditions.

This approach can work well in conjunction with “data from other more established concurrent or retrospective data elicitation measures” (Smith and Renaud 161). Therefore, along with the dynamic trace provided by the eye gaze data, offline measures are additionally used to set the groundwork for a holistic research approach. Vocabulary tests conducted before and/or after the experiments have been previously implemented by many scholars as an offline measure (Montero et al.; Kovacs and Miller; Zheng et al.; Harji et al.).

By and large, adding subtitled text as language input is expected to generate positive vocabulary learning results. Our study’s overarching research question (RQ1) concerns the level of effectiveness of bilingual subtitles (as opposed to monolingual interlingual) on streamed videos in vocabulary learning for young learners of English at an intermediate level. The second research question (RQ2) addressed in the current study is whether the variable of positioning renders different research results. Is the learners’ performance during vocabulary tests linked with the variable of having the native or foreign language subtitles on the top line? Considering these results can help viewers to evaluate the trade-off between putting aside well-established viewership patterns and accepting the beneficial effects of subtitling in foreign language learning.

Methodology

Given the idiosyncratic parameters of the present study, the experiment at hand has been conducted in an attempt to comply as closely as possible with the directives proposed in the most recent position paper on conducting experimental research in audiovisual translation. The team of scholars set the goal of establishing “standardized experimental protocols and frameworks that will allow to conduct scientifically sound, ethical and replicable studies” (Orero et al. 107).

Ethics

Given the fact that the participants belonged to a sensitive population group (children), the current study received ethical consent both from the ethical committee of Aristotle University
and the Greek Ministry of Education. Additionally, the film material used for the experiment was selected from the filmography rated as appropriate for viewing by Greek students.23

Participants

A total number of 79 primary and secondary school students (aged 11-13) volunteered for the study. Participants’ parents gave written, informed consent after reading information on the nature of the study, as well as the exact use of the data. All of them are native speakers of Greek and have lived both in rural and urban environments. The samples were intentionally collected to ensure representability from both types of environments. Prior to the experimental session, placement tests were run by the participants’ instructors to assess their language level objectively. They were found to be of an intermediate language level, which is in accordance with the LLN programmers’ statement of this tool, being primarily geared towards people who have already built essential vocabulary and grammatical knowledge.

Short Memory Test

All participants conducted the Digit Backwards Recall of the Wechsler Adult Intelligence Scale (WAIS) as a measure of verbal working memory (St Clair-Thompson). The term “working memory” is used to describe a more complex system responsible for both the processing and storage of information during cognitive tasks. For the digit recall test, participants were asked to recall sequences of digits spoken aloud by the experimenter in the same order.24 The task was used as a screening tool for participation in the study and 6 participants were excluded from the final dataset due to their performance exceeding beyond the Mean +/-2SD (standard deviation) range per Grade grouping. After the exclusion of the six participants, the remaining ones proceeded with the experiment. All participants (n=73) were native speakers of Greek (43 male 30 female; age range: 11-13). The crossbar chart below describes the demographics of the study. As observed, the combinations of Sex, Grade and Area variables are shown as percentages and absolute values (Fig. 4).

23 A detailed list can be found at edu.klimaka.gr/drasthriothtes/genika/779-egkekrimen-kinematografika-erga-gia-mathites
24 Sequences were constructed randomly and without replacement from the digits 1 to 9 and were presented at the rate of one per second. Following two practice trials, testing began with a maximum of six trials with one digit. If four trials were recalled correctly, the participant was credited with correct recall in all six trials at that list length, and the number of digits was then increased by one. When three or more lists of a particular length were recalled incorrectly, the testing procedure ended. The score was calculated as the number of trials on which the digits were recalled correctly.
Exploring the Educational Potentials of Language Learning with Netflix Tool: An Eye-Tracking Study

Design and Pre-Test Materials

Pilot tests were performed before running the actual experiment in order to adjust parameters and overall duration, and additionally assess the reliability and sensitivity of the tasks. A collectively administered, multiple-choice test was distributed to students in written form ten days before the experiment, and the participants received no feedback on their performance. The particular details of the study were not initially revealed to the participants; the latter were informed on the monitoring process itself and study goals regarding the effectiveness of different subtitling types. A pre-treatment questionnaire was used to establish comparability among students in terms of their familiarity with certain lexical items that were included in the dialogue of the video clip they were going to watch during the main experiment. The items were previously randomized and presented along with several distractor items. A total of ten words were included, five of which were nouns and five were verbs (one item in each category belongs to low-frequency words). Most typical eye-tracking tasks are a part of a mix-method research design, including elements of some off-line tasks that have been used to study language acquisition in children (Sedivy 133).

Test Materials

The video material used for our study was taken from the opening scene of the BBC documentary titled Planet Earth, Episode One: “From Pole to Pole.” As described in the online episode guide, “[t]his episode journeys across the planet, from pole to pole, following the influence of the sun and discovering how its seasonal journey affects the lives of all who live on earth.” The plot structure is rather conventional for a nature documentary with slow-paced

Fig. 4 Cross bar charts for (a) distribution of grade in the area, (b) sex of the participants in each area and (c) sex of the participants in each grade.

25 The selection of this type of lexical items was guided by previous research findings that have shown lack of passive vocabulary acquisition due to the difficulty of the material used in the pre-post vocabulary tests, mainly composed of low-frequency words (Birulés-Muntané and Soto-Faraco).

26 A brief description of the program can be found at www.bbc.co.uk/programmes/b0074sf9
narration accompanied by recorded footage. We presented the same video excerpt (duration 2.53 min) in three different versions. The three experimental conditions are distinguished as follows:

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL_Video Screenshot</td>
<td>(Fig. 5) (the participants watched the movie with bilingual subtitles, Greek subtitles on the top line and English Subtitles on the bottom)</td>
</tr>
<tr>
<td>EN_Video Screenshot</td>
<td>(Fig. 6) (the participants watched the movie with bilingual subtitles, English subtitles on top and Greek Subtitles on the bottom)</td>
</tr>
<tr>
<td>GR_Video Screenshot</td>
<td>(Fig. 7) (the participants watched the movie with standard monolingual Greek subtitles only)</td>
</tr>
</tbody>
</table>

![Fig. 5. EL_Video Screenshot.](image)

![Fig. 6. EN_Video Screenshot.](image)

![Fig. 7. GR_Video Screenshot.](image)

All versions had a slow subtitle presentation rate with almost verbatim translations of the narration. No participant had been previously exposed to this video material. This show was chosen because of its critically acclaimed quality, its educational orientation and the relatively low-speed dialogue and vocabulary density. Captioning may not be suitable for all materials and viewers at all levels of language proficiency. As noted by Danan, beginners/intermediate

---

27 According to its description on IMDB, *Planet Earth* is an Emmy Award-winning series consisting of eleven episodes that took five years in the making. It is the most expensive nature documentary series ever commissioned by the BBC, and the first to be filmed in high definition ([www.imdb.com/title/tt0795176/](http://www.imdb.com/title/tt0795176/)).
students can only benefit from the subtitled material if it is carefully adapted to their level and not far beyond their linguistic competency (71).

Data Collection and Analysis

The eye-tracking experiment was conducted in designated areas inside the schools. For the purposes of this experiment, we used the portable eye-tracker TOBII X2-60. Remote eye-trackers, similar to the one used in the current study, fastened to the bottom edge of the screen and did not restrain the participants from moving their heads. Setting up the eye camera and performing a proper calibration routine is a task of utmost significance (Holmqvist 7-8). Once each calibration process is complete, participants use the computer as they normally would, and their eye movements are captured remotely. The data analysis software, in our case TOBII STUDIO 3.2.0., was then used to create video files in the form of screen capture, recording the location and duration of each participant’s eye fixation superimposed on the screen (Smith and Renaud 153).

The participants were pseudo-randomly assigned to each condition of the experiment balancing out gender representation across the three conditions. They were only included if they declared not having watched the specific TV series before. Each student participated in the data collection session individually. To preserve anonymity, all names were replaced with numbers assigned to their sessions and the reference file was kept in the lab files repository.

Finally, the Areas of Interest (AOIs) were manually marked. For the bilingual videos EL and EN, three areas of interest were drawn namely the image on screen left (Rectangle) and two more for each subtitled line as shown in the chart below (L1-Viewers Native Language/Greek, L2-English). Each area was marked by a different color (Fig.8) and delineated a particular subtitled version.

<table>
<thead>
<tr>
<th>EL_VIDEO ELTOP</th>
<th>L1 SUBTITLED VERSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL_VIDEO EL BOTTOM</td>
<td>L2 SUBTITLED VERSION</td>
</tr>
<tr>
<td>EN_VIDEO ELTOP</td>
<td>L2 SUBTITLED VERSION</td>
</tr>
<tr>
<td>EN_VIDEO EL BOTTOM</td>
<td>L1 SUBTITLED VERSION</td>
</tr>
<tr>
<td>GR_VIDEO</td>
<td>L1 SUBTITLED VERSION ONLY</td>
</tr>
</tbody>
</table>

Fig. 8. Video conditions types.

As indicated below, each bilingual subtitle (EL and EN Videos) is marked by two AOIs, one for each language. In the case of monolingual subtitles (GR_Video) only one area of interest is marked (Fig. 9).
Comparing Attention Distribution over Subtitle Area and the Image between Sexes

Before moving on to analyze and compare the three subtitled versions, we need to examine whether the presence of bilingual subtitles has monopolized the participants’ attention, preventing them from noticing the non-verbal visual cues on screen. Additionally, we investigate whether there has been a statistical difference between male and female participants in their attention distribution. Firstly, the variable of fixation duration was used since it corresponds to the total time or total duration of the learners’ eye gazes at a specific area of interest. To derive meaningful results regarding the fixation duration on the rectangle (screen area without subtitles) and non-rectangle areas of interest (subtitles), the following boxplots were constructed.

As observed here, the median fixation duration is higher on the rectangle than on the non-rectangle areas of interest (Fig. 10). The highest median is detected among female participants attending primary school. Generally, male participants seem to have a low median fixation duration both on rectangle (screen area without subtitles) and non-rectangle videos (subtitled areas) and demonstrate widespread attention distribution. To test whether each of the variables of
fixation duration show a significant difference between male and female participants, the following hypotheses have been stated:

H₀: There is no statistical difference in the average fixation duration between male and female participants.
H₁: There is statistical difference in the average fixation duration between male and female participants.

To determine which of the hypotheses is valid, all variables were tested using t-test based on the sex of the participants. Results showed that hypothesis H₁ can be accepted since male participants have significantly less average fixation duration (μ=22,37 sec) in comparison to female participants (μ=33,61 sec). Between the two participant types, a statistically significant difference is reported at 5% significance level t(23) = -2.177, p = 0.040). This indicates that there is a difference in average fixation duration based on the sex of the participants.

**Comparing Fixation Duration and Fixation Count on Different Video Conditions**

Previous research has shown that subtitles do act as gaze attractors and that “switching attention from the visual image to ‘reading’ the subtitles happens effortlessly and almost automatically” (D’Ydewalle and De Bruycker 196). In line with the relevant literature, this tendency is also depicted in the visualizations of the gaze plots and heat maps captured within the same video. The first frame captures a point with no subtitles on screen and the other two are taken seconds later, at the onset and at the presence of subtitled material. The gradual shift of focus is depicted in the three consecutive film stills shown below both in the form of gaze plot visualizations for the monolingual subtitling condition (Fig. 11) and in the form of heat map visualizations for the bilingual subtitling condition (Fig. 12).

![Fig. 11. Gaze plot visualizations of consecutive film stills (monolingual subtitles).](image-url)
Similarly, visual attention is relatively scattered in each step except for the first stills (Fig. 11 and Fig. 12). In the bilingual subtitles even if there is no presence of subtitles, in the first picture, the “no active subtitle, continue playback” shaded caption embedded in the platform’s interface, as shown below, is obviously attracting part of the participants’ attention (Fig. 13).

To test whether there is a difference in the variables of fixation duration and fixation count among the participants’ fixations on a specific type of subtitled videos, the following hypotheses have been stated:

$H_3$ Participants spend more time looking at the bilingual subtitles compared to standard monolingual ones.

$H_4$ Participants do not spend more time looking at the bilingual subtitles compared to standard monolingual ones.

Focusing on the data in the figures below, the first hypothesis ($H_3$) is validated since both measurements (total fixation duration in seconds and fixation counts) indicate that the participants spend more time fixating on the bilingual subtitled area (Fig. 14) (Total Sum Fixation Duration variable EL_Video n= 39.31 and EN_Video n=47.87) compared to the monolingual subtitled area (Standard n=31.84). Likewise, as regards the variable of fixation count, participants’ record of fixation counts is higher on the bilingual (Total Sum Fixation Count variable Earth_EL n=175.11 and Earth_EN n=196.96) compared to the monolingual subtitled area (Earth_GR n=147.77) (Fig. 15).
Fig. 14. Fixation Duration Mean for the two bilingual subtitled video conditions and the standard monolingual subtitled version.

Paying closer attention to the results drawn only from the two bilingual conditions, the EN video condition proved to be more gaze-attracting towards the subtitled area compared to the EL video version. We further hold that participants were inclined to follow their native language version subtitle line irrespective of its positioning. The following chart depicts the distribution of the participants’ total fixation duration in the two bilingual subtitled conditions. It clearly demonstrates that the viewers’ attention was directed towards the EL top and EN bottom (light and dark green respectively) areas of interest (Fig. 16).

Fig. 15. Total Fixation Count Mean for the two bilingual subtitled video conditions and the standard monolingual subtitled version.

Fig. 16. Total Fixation Duration Mean for the two bilingual subtitled video conditions.
The Post-test Analysis

The final step for the completion of the current study involved a multiple-choice vocabulary test that was administered directly after each viewing session in pen-and-paper format. All the lexical items appeared once during the extract. As explained above, ten days before the experiment, the same test was given to the participants who subsequently repeated it after the participants’ exposure to the material. The vocabulary test consisted of the same definition-matching items. Analyzing whether the answers have changed after the video viewing can yield interesting results regarding the possible passive acquisition of new words.

Regarding vocabulary acquisition, our results showed pre-post improvements in the bilingual conditions, though the most significant results were found in the EL_Video condition (Greek subtitles on the top line) according to the calculation based on the grand total. By analyzing all the words-variables for each video type, then in total and as supported by the following bar chart (Fig.17), the EL_Video condition attracted a total 27% of post-test correct answers (namely 18% pre-test wrong – post-test correct and 9% pre-test correct – post-test correct). The sum of the corresponding post-correct categories for the EN_Video comprised 23% (namely 17% pre-test wrong – post-test correct, and 6% pre-test correct – post-test correct) and, in the case of monolingual GR_Video, the post-test corrections amount to 19%. According to the following graph, EL_Video condition has improved students’ performance in vocabulary acquisition in most of the cases, compared to other video types.

![Bar Chart](image)

Fig. 17. Frequency analysis of the test results according to each video type.

Discussion

Using audio-visual material as a teaching resource is a common practice in foreign language learning. It contextualizes the process, facilitating learner exposure to authentic input and real-life conditions. Its flexibility in terms of instructional settings is also highly appreciated since it can be used in different circumstances from the physical classroom to online programs. Its versatility in terms of the wide range of available themes and content further enhances the expansion of its use in the learning process. The great value of video, as Talavan puts it, lies in
its “combination of sounds, images, and sometimes text (in the form of subtitles), together with the socio-cultural information about habits, traditions and culture” (5).

One of the questions that has guided this research concerns the criteria for the selection of the most appropriate subtitled video material according to the learners’ age and knowledge of the foreign language. With the aid of eye-tracking technology, useful insights have been gained concerning what learners attend to or notice while watching a new type of subtitling practice that has emerged during the last years in the new media landscape.

The first finding to emerge from this study was the fact that, in the case of this novel type of viewing experience with the presence of the bilingual subtitles, the audience still focuses a significant share of their attention on the area of the film action (rectangle area of interest). We can thus infer that the power of non-verbal imagery is still the most dominant factor, enabling viewers to appreciate the audiovisual product as a holistic experience. As far as specific audience types are concerned, female participants of the younger age group (Primary Education) seem to have more stable visual attention allocation while watching audiovisual material. Although there is significant statistical difference between the two sexes, the focus span of male participants still rendered improved performance results in the vocabulary tests that followed the experiment.

Empirical data have also addressed the research question set regarding the learners’ fixations on bilingual compared to monolingual subtitles. When we focus on the juxtaposition of different subtitles types, in accordance with previous research data analysis, the eye-movement data collected in our study show that the participants spent more time looking at bilingual subtitles than monolingual ones. Liao attempts to explain this finding in relation to the appearance of non-native language in the subtitling area (87), a condition that holds true in our case study as well since, as far as Greece is concerned, watching non-native language subtitles is not a common practice for children of this age. Given the fact that the original spoken language is the same in all video conditions, the higher fixation scores in the bilingual-subtitled areas could be explained as an extra effort on the viewers’ part to process the information or, simply, as an attractive feature that lies outside familiar viewing patterns. Given the novelty of the eye-tracking methodology in similar studies, no conclusive answers can yet be given to explain the participants’ eye-fixations. Comparing the two bilingual conditions, we find that the EN_Video condition received the highest scores in terms of both fixation duration and count. This can be attributed to the presence of intralingual English language subtitles on the top line, an element that students are least familiar with as far as their regular habit formation of viewership is concerned.

With regard to the second research question, that is whether the positioning of each language version plays a role in the learners’ performance, results revealed a relationship between a specific subtitle positioning in the case of bilingual subtitling and post-test success. As evidenced in the statistical data, presented above, the EL_Video condition, whereby native language subtitles were provided on the top line alongside the transcription of the spoken language text on the bottom line has yielded the highest scores in terms of the viewers’ performance.
Limitations and Prospects for Further Research

Limitations

We need to outline certain practical issues concerning the design of the eye-tracking experiment and the analysis of data. Researchers should expect considerable data loss, since the activity levels of young children could potentially jeopardize the process. Children are less tolerant of long experimental sessions (Sedivy 133) and must therefore be constantly instructed to follow the rules of conduct. Furthermore, when sample collection takes place outside the lab, especially in a noisy place (such as an educational institution) special care must be taken so as to minimize any surrounding distractions. Administrative issues regarding the entrance of visitors in school units or forms of research approval can certainly be time-consuming and cumbersome.

In terms of research methodology and data analysis, many issues are not pertinent when working with children and adjustments need to be made. Additionally, although eye-tracking data can be revealing and precise, researchers still need to make inferences, trying to explain the reasons behind the participants’ attention focus. Having said that, any conclusions need to be confirmed given the suggestive nature of the findings discussed in our study. The results would be more firmly supported had there been more tokens demonstrating the relationship between eye fixations and the learners’ performance.

Prospects for further research

Drawing on empirical data, this study has ventured to shed some light on how this type of multimodal processing influences vocabulary acquisition. Eye-tracking research is still in its infancy as far as its implementation on audiovisual translation and foreign language learning is concerned, so new research directions may involve updating and expanding on previous research based on larger participant samples with diverse age population characteristics. Furthermore, viewers may respond differently if experiments are conducted under natural viewing conditions, such as watching an entire feature-film rather than just a small extract or watching a clip in the comfort and familiarity of their homes.

Additionally, implementing this research methodology in different video genres, such as fast action films, news reports or even cartoons could yield potentially diversified results, since the complexity of the visual stimuli in terms of appearance or editing can play a decisive role in luring the participants’ attention. Finally, the next research step is linked with the combination of different methodologies that let us further monitor the reception of audiovisual material, such as measuring participants’ electrodermal activity (EDA) or electroencephalography (EEG). Audiovisual translation is a type of service and apart from the process itself, the way end-users receive its products should be further researched.

To conclude, keeping pace with the constantly evolving reality and the interdisciplinary convergences within the emergent new media culture, new correlations have emerged across the education, the entertainment industry and reception studies. The abundance of choices in an
online participatory environment is here to stay and its complex workings should be further exploited.

This research is carried out/funded in the context of the project “Implementing Audiovisual Media in Education: Evaluation and Application of Eye-tracking Data” (MIS 5005088) under the call for proposals “Supporting Researchers with Emphasis on New Researchers” (EDULLL 34). The project is co-financed by Greece and the European Union (European Social Fund- ESF) by the Operational Programme “Human Resources Development, Education and Lifelong Learning 2014-2020.”

Works Cited


"From Pole to Pole." Planet Earth, directed by Alastair Fothergill, *BBC*, 2006.


O’Sullivan, Carol, and Jean-François Cornu. “History of Audiovisual Translation.” *The


**List of Figures**

Fig. 1. Dual Sub screenshots for different layouts.
(bonigarcia.github.io/dualsub/index.html) .......................................................... 6

Fig. 2. Fleex player screenshot (vocabulary lists) (fleex.tv/en/Home/Index4) ........................................ 6

Fig. 3. Number of Netflix subscribers USA and International in millions ........................................ 6

Fig. 4. Cross bar charts for (a) distribution of grade in the area, (b) sex of the participants in each area and (c) sex of the participants in each grade. .................................................. 11

Fig. 5. EL_Video Screenshot ........................................................................................................ 12

Fig. 6. EN_Video Screenshot ........................................................................................................ 12

Fig. 7. GR_Video Screenshot ........................................................................................................ 12

Fig. 8. Video conditions types ...................................................................................................... 14

Fig. 9. Areas of Interest for bilingual and monolingual videos .................................................. 14

Fig. 10. Boxplots regarding the sum of fixation duration to rectangle (a) and non-rectangle (b) areas of interest .................................................................................................................. 14

Fig. 11. Gaze plot visualizations of consecutive film stills (monolingual subtitles) ..................... 15

Fig. 12. Heat map visualizations of consecutive film stills (bilingual subtitles) .......................... 16

Fig. 13. No active subtitle caption ................................................................................................ 16

Fig. 14. Fixation Duration Mean for the two bilingual subtitled video conditions and the standard monolingual subtitled version .................................................................................. 17

Fig. 15. Total Fixation Count Mean for the two bilingual subtitled video conditions and the standard monolingual subtitled version .................................................................................. 17

Fig. 16. Total Fixation Duration Mean for the two bilingual subtitled video conditions ............. 17

Fig. 17. Frequency analysis of the test results according to each video type ............................... 18