

Captions and Subtitles in EFL Learning: an investigative study in a comprehensive computer environment

FRANCESCA BIANCHI*
TIZIANA CIABATTONI**

0. ABSTRACT

This study is a broad-range investigation into short- and long-term effects of captioning and subtitling in beginner, intermediate, and advanced Italian adult learners of English. Several issues are taken into consideration including content comprehension, vocabulary acquisition, language-in-use, and semantic match between audio and video inputs. All the variables involved were controlled in a single computerised setting. The current experiment partially supports the findings described in the relevant literature. A few discrepancies emerged with some previous studies, but they are probably explained by the different type of material and testing procedure adopted.

1. INTRODUCTION

Ample research has been carried out on captioning (also called “bimodal input” or “L2 subtitled video”), i.e. the display of transcriptions of the utterances of a video, and its effects on L2/FL learning (see for example Baltova, 1999; Chung, 1999; Garza, 1991; Guillory, 1998; Markham, 1989; Neuman & Koskinen, 1992; Price, 1983; Vanderplank, 1988, 1990, 1993). Many of these experiments compared captioned video to audio input only and focused on general comprehension. Price (1983), for example, found that captions significantly improved performance on comprehension regardless of language background. Similar results were obtained also by Markham (1989). The former studied the effects of captioned TV

upon listening comprehension in students of different proficiency levels. The subjects were shown a video with captions and a video with no text aids, then their general comprehension was tested with multiple-choice questions; all the three groups performed significantly better with the captioned video. However, Guillory (1998) noticed that the impact of captioning on learning depends on the gap between student's proficiency level and difficulty of the spoken text: captions cannot compensate for an excessively wide gap.

Captions' impact on vocabulary learning was assessed by Garza (1991), who found that captions increased comprehension and language memorisation in advanced FL learners. Similar results were reported by Neuman & Koskinen (1992): in an experiment with advanced EFL students, those who were shown captioned video had better results in vocabulary recognition and acquisition exercises. Subsequently, Baltova (1999) reported positive effects of captions on content and vocabulary learning also on relatively inexperienced students (grade-11 core French students in Canada) both at short- and long-term level.

The display of a translation of the utterances of a video in a different language is another widely studied phenomenon that has attracted the attention of researchers from two different perspectives: L1 subtitles of L2 aural input (subtitling) and L2 subtitles of L1 aural input (reversed subtitling). Pioneering experiments were carried out in 1981 by Lambert, Bowler & Sidoti (cited in Holobow, Lambert & Sayegh, 1984, and in Danan, 1992) and in 1984 by Holobow, Lambert & Sayegh. They compared different combinations of audio and video monolingual or bilingual input and found that the most favourable condition was reversed subtitling, followed by captioning, monomodal input, and "ordinary" subtitling (in this order). In both experiments, the subjects were 5-6 grade English-speaking pupils who had taken part in a "French immersion" program starting at kindergarten. The pupils tended to «rate themselves as *slightly* more English dominant than French in writing, reading and understanding and *somewhat* more English dominant for speaking» (Holobow, Lambert & Sayegh, 1984: 61). Despite the fact that these two experiments stemmed from «a practical interest in making better use of radio and television in education» (ibid.: 59), the material and procedure adopted included only aural input (a teacher reading a text) and written visual input (a script of the text). These conditions are very far from the use of captions or subtitles in TV programs and movies.

A study in a setting that was closer to actual TV or movie watching, and fairly similar to the setting of our experiment, was carried out by Danan in 1992. In her 3 subsequent experiments, focus was on vocabulary, and the subjects were college students with not very high proficiency in French (30 students in one case, 57 in the second case, and 15 in the third; mixed levels). All experiments used the same 5-minute extract from a French video for learning purposes; in the first experiment the following three conditions were tested: subtitling in English, reversed subtitling, and French audio only; in the second, subtitling was replaced with captioning; in the third, only reversed subtitling and captioning were tested. The video extract was shown twice and the students were tested immediately after the second view on their ability to recall the correct French names of items that were foregrounded by a «clear link with a video image» (Danan 1992: 509) in the video. In the test, the students were guided by the original script with gaps

and an image-only presentation of the video (no sound, and no titles). Before watching the experimental video, however, the students had been given a summary of the scene. In the same study, an attempt was also made to assess long-term effects of bimodal input via translation. However, the interval between the short-term test and the long-term one is not declared in the paper. Their data showed that reversed subtitling, immediately followed by captioning, produced the most favourable results in both short- and long-term measures in beginners as well as higher-level students; however, in the second experiment (the one with the highest number of subjects), the difference between the reversed subtitling condition and the captioning condition was not significant. On the other hand, “ordinary” subtitling (assessed only in experiment 1) seemed to be the least favourable condition, as it led to results that were slightly lower to those of the L2 audio-video only condition. Contrary to Danan’s (1992) results with respect to “ordinary” subtitling, however, Koolstra & Beentjes (1999) found that children exposed to subtitled video acquired a higher number of new words in the foreign language than those who watched the same video with no text aids. They also noted that older children performed better than younger ones, but this was due to their being more frequently exposed to subtitles when watching TV.

As far as text aids are concerned, therefore, ample evidence exists that captions help comprehension and vocabulary memorisation at all levels of proficiency. “Ordinary” subtitling also seems to play some facilitatory role in language acquisition, but the extent of this role is debated. However, these are not the only variables at play when watching videos. Other fundamental elements are kinesic behaviour in the video and semantic match between image and sound.

Kinesic behaviour and non verbal communication play a fundamental role in listening comprehension. Grimes (1990) found that a high degree of correspondence and semantic match between the audio and video channels favoured attention and memory of video texts in L1 subjects. The absence of said semantic match, however, negatively impacted on both faculties. Baltova’s (1994) study indicates that scenes where dialogues were backed up by action or body language tended to be more easily understood by FL students than scenes with static images and unrelated audio. Duquette and Painchaud (1996) investigated the impact of images on L2 vocabulary learning. Their study was carried out on two groups of L2 students: both groups listened to the same tape, but one of them also watched a video showing the actions of what was being described on the tape (high semantic match). Both groups recorded similar overall vocabulary results, but while the audio-only subjects tended to retain primarily higher-frequency words or words that sounded similar to their original language, the video subjects retained other types of words.

Finally, as Vanderplank’s (1988, 1990, 1993) experiments highlighted, taking advantage of text aids in a tri-channel environment requires some kind of strategic adjustment. Some of the students he worked with, in fact, declared feeling initially disturbed by subtitles, but they eventually managed to develop adequate personal strategies to process the three channels. He also noted that such strategies were more readily present in students coming from countries where subtitling is a common occurrence.

The experiments on text aids reported above were carried out each on a different type of video material, spanning from educational videos, to television in-

formative programmes, real-video, and films, on subjects of different ages, and with very different procedures. If on the one hand this seems to enhance the general validity of the findings, on the other hand it makes it difficult to compare results at a detailed level and almost impossible to analyse trends in terms of image-audio-text relations. Finally, in most cases research has focused on short-term effects of text aids, rather than long-term ones, and in 2004 Danan still advocated the systematic collection of long-term data.

The current study attempted to investigate the role of captioning and subtitling in an experiment where all the different variables involved were controlled in a single setting. The following variables were identified and controlled: short- vs. long-term effects of captioning and subtitling on content comprehension, vocabulary acquisition, and language-in-use issues; students' level of proficiency; semantic match between the audio and video inputs. In particular, this study aimed to provide insight into the following issues: if we consider different types of semantic match between audio and video inputs, which type of text aid proves more useful?, with respect to which type of language feature?, and for which level of proficiency?

2. METHOD

The experiment was developed at the University of Pavia (Italy) within a course of English for the faculty of psychology and targeted psychology students; participation in the experiment was voluntary, but allowed for a small reduction in the English exam workload.

2.1 PARTICIPANTS

A total of 107 students volunteered for the study. After initial assessment of the subject's proficiency in English, and in an attempt to create comparable groups whose composition could mirror the distribution of the total population taking part in the experiment, beginner, intermediate and advanced students were separately and randomly assigned to one of three groups: experimental group 1 (EG1), with captions; experimental group 2 (EG2), with subtitles; and a control group (CG), with no text aids. Reversed subtitling was not considered for two main reasons: 1. it is not a usual condition in film watching for autonomous-learning purposes; 2. according to Holobow, Lambert & Sayegh's (1984) results, this condition was highly comparable to captioning.

Unfortunately, due to either personal or technical problems, some students did not have the chance to complete all the phases of the experiment and their data could not be included in the final database. Therefore, the population for this experiment eventually comprised a total of 85 volunteer adult participants in the 18-45 age range. In terms of knowledge of the English language (assessed at the very beginning of the experiment), 17 subjects could be considered beginners, 45 intermediate learners, and 23 advanced learners. Only 13% of the subjects were males, but their distribution was balanced across language levels (5 beginner, 4 intermediate, and 4 advanced learners of English).

	EXPERIMENTAL GROUP 1 CAPTIONS (ENGLISH)	EXPERIMENTAL GROUP 2 SUBTITLES (ITALIAN)	CONTROL GROUP (NOTHING)	TOTAL
Beginners	5 (21%)	7 (21.2%)	5 (18%)	17 (20%)
Intermediate	12 (50%)	18 (54.5%)	15 (54%)	45 (53%)
Advanced	7 (29%)	8 (24.3%)	8 (28%)	23 (27%)
Total	24 (100%)	33 (100%)	28 (100%)	85 (100%)

Table 1. Distribution of beginner, intermediate and advanced students within the experimental and control groups.

As Table 1 shows, despite the attempts to create perfectly balanced groups, the fact that 22 participants had to be excluded from the final database determined differences in numbers between the three groups, but the composition of each group still mirrored the composition of the total population taking part in the experiment in terms of proficiency in English.

According to the data gathered at the beginning of the experiment, none of the students had watched the film from which the first scene was taken, nor had they read the book that inspired the film; on the other hand the film from which the second scene was taken was known to a few students ($N = 12$), but they had watched it more than 10 years before.

2.2 EXPERIMENT OUTLINE

The experiment was organised along the following three phases:

PHASE ONE: PRE-TEST. A collectively-administered test in written form composed of four tasks. The pre-test aimed to assess the participants' level of English before the beginning of the experiment, as well as their knowledge of the words, phrases, and linguistic phenomena targeted in Phase Two. The participants were given a maximum time span of one hour to complete the pre-test (four tasks in all). Before distributing the test papers, the researchers briefly explained the general aim of this first phase (assessing student's proficiency) with reference to the entire experiment, and tried to motivate the subjects towards a correct and honest accomplishment of the tasks.

On the basis of pre-test results, each student was assigned to one of the two experimental groups or to the control group, in an attempt to create balanced and comparable groups.

PHASE TWO: COMPUTERISED VIDEO TEST. This test was carried out on an individual basis, with the aid of a specially developed computer application. On individual computers with headphones, the subjects watched a series of clips from two famous films in English, accompanied by captions, subtitles in Italian, or nothing, according to the group each subject had been assigned to. At the end of each clip, a series of multiple-choice questions was presented to test the subject's comprehension in terms of content, vocabulary, and use of lexico-grammatical phrases; at the end of each series of questions the subjects had the possibility to watch

the entire film clip again and then review their answers up to two times. This mechanism allowed maximum freedom to the subjects, who could work at their own pace and view the clips one or more times according to their habits, level of interest, and commitment to the task. This phase, in fact, was intended to simulate a scenario of an adult intentionally watching a film as a means of learning English. In such a scenario some more motivated and systematic learners would go over the same scene more than once if they felt they had not grasped or understood one or more words or utterances; other types of learners, instead, tended to be content with understanding the general meaning of scenes on the basis of a few keywords and the accompanying pictures and would not bother to watch the same scene twice, as this is a time-consuming task that delays the development of the plot. In our application, re-watching a scene was not at all compulsory and was rather time consuming; therefore, subjects who would not go over the same scene in the real scenario would presumably not do it in our simulation. However, it must be noted that this simulated scenario included two features that should facilitate learning: short video segments and criterion-based questions (Canning-Wilson, 2000). Phase Two took place no more than seven days after Phase One.

PHASE THREE: POST-TEST. A repetition of Phase One. The participants were collectively administered the same written exercises that were given in the pre-test, following the same procedure. This phase took place one week after Phase Two and aimed to assess the long-term effects of captions and subtitles on language learning.

2.3 MATERIAL

In Phases One and Three, the students were administered four written multiple-choice tasks in pen-and-paper format. Task One was a multiple-choice cloze test on grammar, with items of increasing difficulty focusing on verb tense usage, modal and auxiliary verbs, pronouns, comparatives and superlatives, and prepositions. Each item was composed of a single self-contained sentence in English with a gap, accompanied by four possible solutions for the gap. This test had been developed and used for years as a placement test in a local private school of foreign languages. Task Two aimed to test the students' general lexical knowledge in English. The students were given a list of words and asked to circle the correct synonym among the four alternatives that appeared to the right of each word. The test, which follows the structure and logic of the PMA 11/17 test (Thurstone & Thurstone, 1981), a standard lexical test, had been originally developed and used by Palladino and Bianchi to assess lexical abilities in adult learners of English (Palladino & Bianchi, forthcoming). The results of Tasks One and Two, taken together, were used to assign each subject to the beginner, intermediate or advanced group (scores <28, 28-43, and >43 respectively).

Tasks Three and Four were structured so that they could be compared to the results obtained by the students at the computer. Task Three targeted vocabulary and resembled Task Two in form, but the words were chosen among those used in the film clips on the basis of their prominence in the dialogues and relevance for the comprehension of the clips. Task Four focused on the pragmatic use of lexico-grammatical phrases taken from the film clips chosen for the experiment.

This task will be referred to as “language-in-use” and was composed of multiple-choice questions referring either to scenes from *Harry Potter* or *Fantasia*. For each item, four possible answers were provided. Some items asked the students to decide on the use of phrases such as “how about a film”, “you’d make a good tennis-player”, either by using the given phrases to complete sentences or by choosing a correct pragmatic description (such as “statement”, “question”, “order”, “exhortation”). The other items asked the meaning of phraseological or idiomatic expressions (“drop the other shoe”, and “what’s going on?”), tested the use of prepositions, or asked about the circumstances for the use of the genitive noun phrase structure. The mixed nature of the exercises was a direct consequence of the dialogues in the chosen film clips, which were fairly simple and repetitive in terms of grammatical features.

Phase Two was entirely computerized. The creation and administration of the material was constrained by a series of needs and considerations. Phase Two intended to simulate a real home-video scenario where a student watches a film on DVD and takes advantage of the text aids provided (captions or subtitles). In a real context, images are displayed full-screen and with high resolution, and the audio and the texts are perfectly synchronised. Furthermore, the student can view the same scene more than once, if s/he wants to. Finally, there was the need to assess the student’s comprehension by means of a high number of questions and quantitative analysis of the answers. To achieve all this on a computer, a program, called V.A.L. (View And Learn), and a dedicated application, called CA.S.T.ing (Caption and Subtitle Test-ing), were created. V.A.L. allows the seamless integration of audio, video, hypertext, and text files. It can be used as a research tool to test, for example, teaching methods, or as a tool for the creation of highly interactive multimedia applications for individual, self-paced learning of a foreign language or any other subject-matter. It includes a multiple-choice and limited-answer testing system, as well as database and statistic analysis functions for an automatic assessment of the students’ performances. CA.S.T.ing is an application of V.A.L. that was created specifically for the current project and offers the following features: full-screen, high-resolution video; synchronised audio; well-visible, and synchronised original text; possibility to select the text (captions, subtitles, or nothing) at the very beginning of the session; audio control commands; possibility to re-play the same scene more than once; alternation of film clips and multiple-choice questions; preliminary window for gathering general information about the students; automatic recording of the students’ answers in a database; automatic recording of the length of each session.

Therefore, CA.S.T.ing included selected clips from two films: *Fantasia* (Walt Disney)¹ and *Harry Potter and the philosopher’s stone* (Warner Bros)². The scenes were chosen according to the following criteria; (a) each scene is self-contained and fully understandable even when detached from the rest of the film; (b) the scenes clearly differ in terms of event-word-image relations: while in *Fantasia* the images, although matching the content of the text, do not help one understand what is said, neither at a linguistic nor at a cognitive level, in *Harry Potter* the pictures are almost fundamental to an understanding of the meaning of words (e.g. the names of the different kinds of quidditch balls) and sentences (e.g. the game commentary). The film clips were presented in English, accompanied by captions, subtitles in Italian, or nothing, according to the group each partici-

pant had been assigned to. At the end of each clip a series of multiple-choice questions was automatically displayed to test the subject's comprehension in terms of content, vocabulary, and use of lexico-grammatical phrases; at the end of each series of questions the subjects had the possibility to watch the entire film clip again and then review their answers up to two times. General information about the participants (such as age and mother tongue), and whether they had watched the two films before, were also automatically collected at the beginning of this phase.

The experiment was carried out at the very beginning of the academic year and stretched over a total of three weeks. Its start coincided with the beginning of English lessons at the faculty of psychology; however, given the scant number of hours of English the students were exposed to during that period (four hours in all) and the specialised content of the course, it is highly improbable that the academic English lessons influenced the results of the experiment. The academic lessons, in fact, focused exclusively on psychology research articles, a written genre characterised, like most other academic types of written texts, by highly specialised lexicon, absence of idiomatic expressions and colloquialisms, and prevalence of passive and infinitive constructions. Furthermore, the first few lessons were taught in Italian, as they simply aimed to provide the students with basic general information about this particular genre. On the other hand, the experiment tested comprehension and acquisition of general vocabulary, colloquial and idiomatic expressions, and use of phraseology in informal spoken contexts.

3. RESULTS AND DISCUSSION

All the analyses were carried out on mean scores, standardised according to the following parameters: number of subjects per group, number of items per task, and subject's proficiency level. Analysis of immediate comprehension was based on mean results obtained in Phase Two. Long-term acquisition was measured on difference scores (% DELTA) calculated comparing Phase Three mean results with Phase One mean results. A direct comparison between Phase Two and Phase One/Three tasks was impossible, given the different structural and methodological features characterising the three phases (electronic format and the possibility to look for the correct answers by watching the film clips up to two extra times in one case; pen-and-paper format and no reference text for the answers in the other cases).

3.1. IMMEDIATE COMPREHENSION

The data gathered in Phase Two made it possible to evaluate the impact of captions and subtitles in the immediate comprehension of content, vocabulary, and use of lexico-grammatical phrases. The findings will be presented according to task, with details regarding students' proficiency level, and type of film.

3.1.1 Content

As Figure 1 shows, at beginners' level, EG2 participants (with subtitles) fared better in the comprehension test than EG1 (with captions) and control participants,

in both types of films. However, while the difference between the three groups was rather marked when considering the questions referring to *Fantasia*, in the case of *Harry Potter*, EG1 and CG's comprehension answers were not significantly worse than EG2's, with the control subgroup faring slightly better than EG1. Furthermore, comprehension was generally higher when watching *Harry Potter* clips; an indication that the students' comprehension was greatly helped by the images.

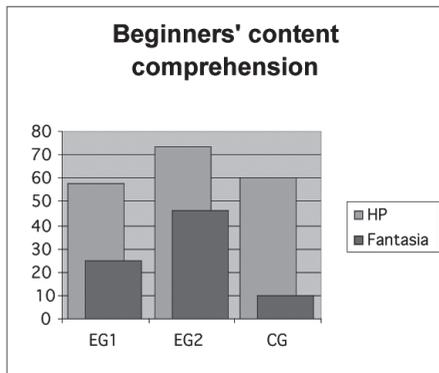


Figure 1. Short-term results: beginners' mean scores in the content comprehension test.

With regard to intermediate students, it seems that content comprehension (Figure 2) was favoured by Subtitles, and this is particularly evident in the questions regarding *Fantasia*. In the case of the *Harry Potter* clips, the results obtained by EG1 and EG2 intermediate students were almost identical, the EG2 sub-group having fared only 0.4% better than the EG1 sub-group.

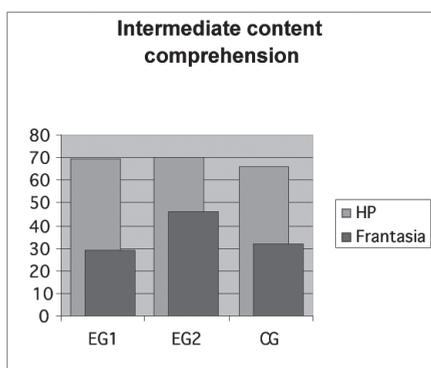


Figure 2. Short-term results: intermediate students' mean scores in the content comprehension test.

Interestingly enough, a direct comparison between EG1 and CG intermediate students shows different results with respect to the two different types of film: the EG1 sub-group fared better than the CG sub-group in the questions on *Harry Potter*, but worse in those on *Fantasia*, with an opposite trend to that of the beginner participants.

Advanced students' results in content comprehension highlighted the same trend with both types of film, with the EG2 sub-group scoring slightly higher than the EG1 one and significantly better than the CG sub-group (Figure 3).

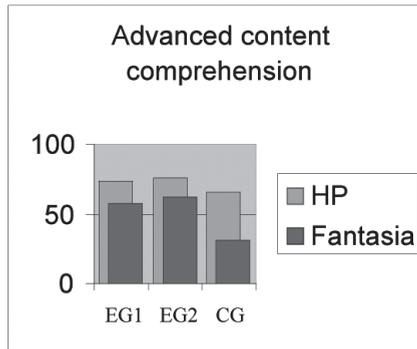


Figure 3. Short-term results: advanced students' mean scores in the content comprehension test.

To sum up, in the content comprehension tasks EG2 students (with subtitles) obtained the best results, regardless of their proficiency level, and of the type of film. This result is expected given that subtitling is processed automatically and content comprehension can logically be facilitated by text in the mother tongue. On the other hand, captions proved more useful than no-text input for beginners and advanced students, which is in line with previous literature (Markham, 1989). The same was not true, however, for intermediate students. Finally, when semantic match was high (*Harry Potter* clips), content comprehension was constantly higher regardless of proficiency level and type of visual aid, and differences between experimental and control groups were less marked. This result is clearly in line with the literature and supports the fundamental role of images in general content comprehension (Baltova, 1994).

3.1.2 Vocabulary

When it comes to vocabulary comprehension (Figure 4), the best results at beginners' level were obtained by the control group; however, when text was displayed on screen, subtitles were of greater help than captions. The trend was identical for both types of films, with slightly higher scores in the case of *Harry Potter*.

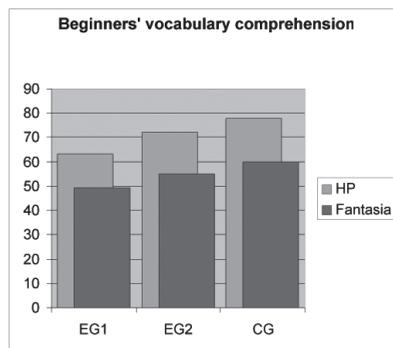


Figure 4. Short-term results: beginners' mean scores in the vocabulary comprehension test.

At intermediate level (Figure 5), both experimental and control groups obtained good results in *Harry Potter*, with a slight advantage for EG1. With regard to *Fantasia*, EG2 emerged as the best group, with a higher score by 8%. Interestingly, the profiles of intermediate students with respect to vocabulary are similar to the intermediate profiles in the comprehension tests, except for a smaller difference in scores between *Harry Potter* and *Fantasia*. The same is not true for the other two proficiency levels.

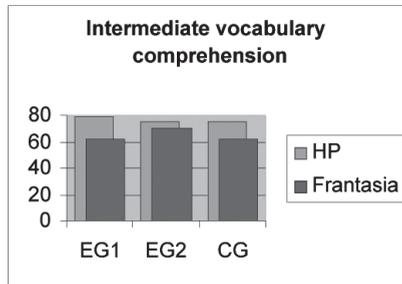


Figure 5. Short-term results: intermediate students' mean scores in the vocabulary comprehension test.

At advanced level, no significant trends can be seen, as the three groups' results with each film were almost identical (Figure 6).

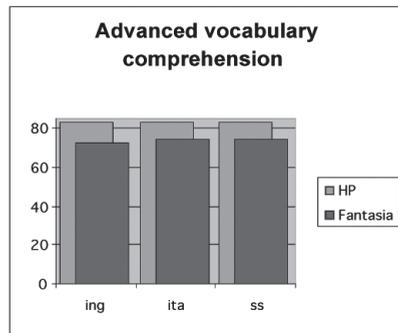


Figure 6. Short-term results: advanced students' mean scores in the vocabulary comprehension test.

The profiles of the three proficiency levels have very little in common, except for higher scores when semantic match among the different communication channels was higher (*Harry Potter*). A comparison between EG1 and EG2 students across proficiency levels (Figures 7 and 8) seems to show that captions were less useful for vocabulary comprehension than subtitles, especially when proficiency was lower or images did not particularly assist dialogue and plot comprehension.

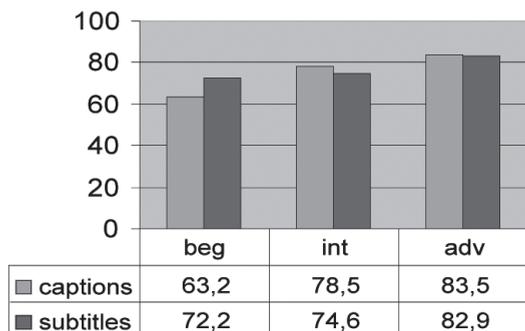


Figure 7. Short-term vocabulary results: comparison between EG1 and EG2 in *Harry Potter*.



Figure 8. Short-term vocabulary results: comparison between EG1 and EG2 in *Fantasia*.

This contrasts with Danan's (1992) results only partially, as she tested vocabulary under conditions of high semantic match only. In our experiment, vocabulary results with *Harry Potter* clips were closer to Danan's, at least as far as intermediate and advanced students were concerned. Furthermore, different testing techniques were adopted in the two experiments: Danan tested vocabulary by giving the students a gapped version of the script, while in the current experiment the participants were asked to select the correct synonym in a multiple-choice exercise, a testing procedure that was closer to the one adopted by Koolstra & Beentjes (1999).

3.1.3 Language-in-use

Beginners' results in the language-in-use questions (Figure 9) showed a similar trend to beginners' vocabulary results, in that EG1 scored worse than EG2, which in turn scored worse than CG, in both types of films. Slightly higher mean scores were recorded with questions on *Harry Potter* in the experimental groups, but not in the control group.

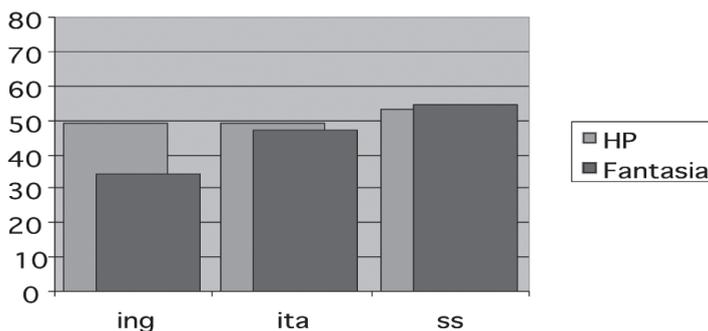


Figure 9. Short-term results: beginners' mean scores in the language-in-use test.

In terms of language-in-use comprehension, the results of the intermediate students showed no significant differences with reference to *Harry Potter* clips, with a slight trend towards an increase from captions to subtitles to no-text-aid. This trend is similar to the beginners' trend, although less pronounced. Fairly different was the trend with questions on *Fantasia*, where the control group scored slightly higher than EG1, and EG2 came last (Figure 10).

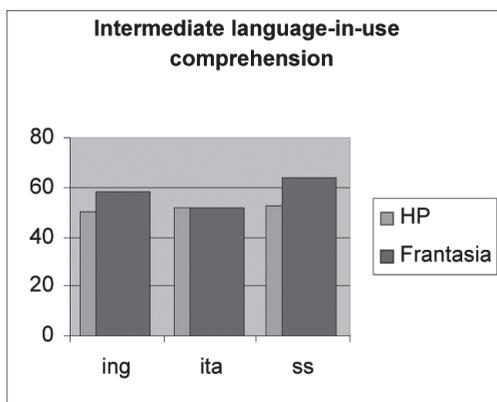


Figure 10. Short-term results: intermediate students' mean scores in the language-in-use test.

Finally, in the language-in-use tasks, EG1 advanced students obtained the highest scores with both types of film. However, while differences are not significant with *Fantasia* clips, with *Harry Potter* clips the control sub-group scored the worst results (Figure 11).

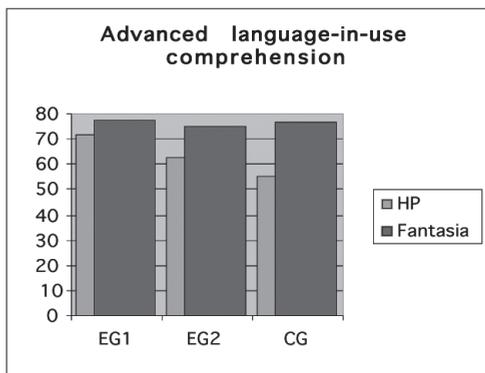


Figure 11. Short-term results: advanced students' mean scores in the language-in-use test.

Plotting beginner, intermediate, and advanced student data without taking into consideration the difference between the two types of film highlighted an interesting general trend along the proficiency line, which sees a gradual passage from text aids in general and captions in particular limiting comprehension in lower proficiency groups to the complete opposite with advanced students (Figure 12).

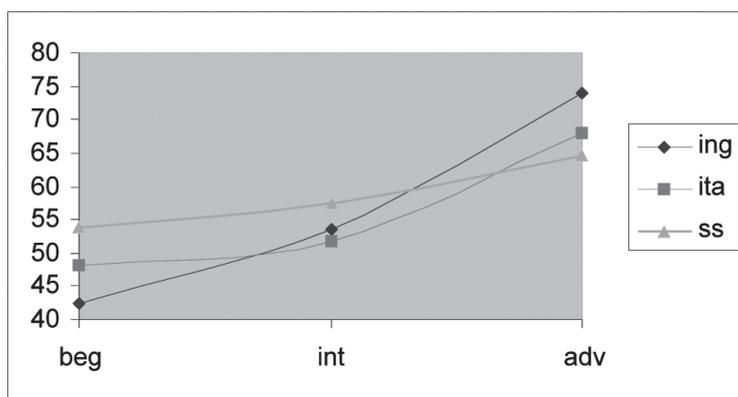


Figure 12. Short-term results in the language-in-use task with reference to text aids.

Finally, a comparison between beginner, intermediate, and advanced student mean results in the language-in-use task in the two types of film regardless of the presence of textual aids offered an unexpected perspective on the role and impact of different types of images (Figure 13). In fact, while beginners obtained generally higher results with *Harry Potter* (+5.3%), intermediate and advanced participants obtained higher scores with *Fantasia* (+6.3 and +13.3 respectively).

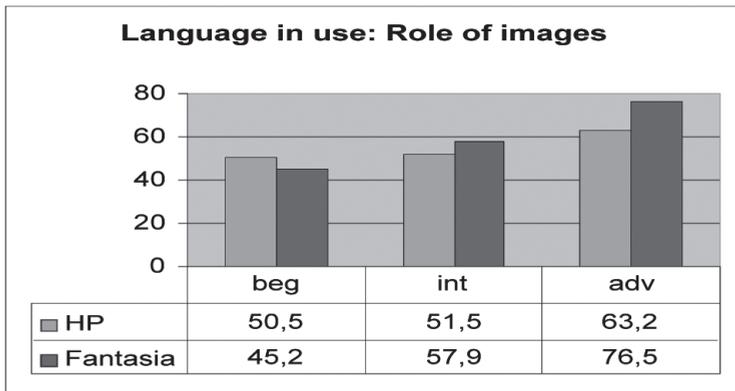


Figure 13. Short-term results in the language-in-use task with reference to film type.

3.2. LONG-TERM ACQUISITION

Long-term acquisition was measured on mean difference scores (% DELTA) calculated comparing Phase Three mean results with Phase One mean results, task by task. Specific vocabulary (Task Three) and language-in-use (Task Four) results were first analysed comparing EG1, EG2 and CG; then other parameters such as participant's level and type of film were taken into consideration.

3.2.1 Vocabulary

A comparison between EG1, EG2 and CG results regardless of proficiency differences (Figure 14) showed that text aids can be useful to learn vocabulary, a finding that is in line with Paivio's (1986) dual coding theory and what described in the reported literature on captioning. In particular, and in contrast with Danan's (1992) results, subtitles seemed to be slightly more fruitful than captions, generally speaking.

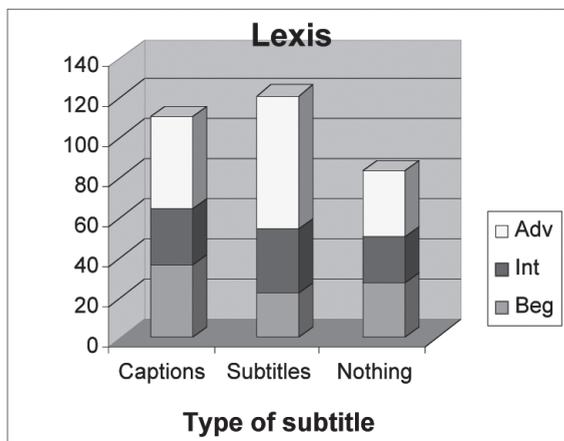


Figure 14. Long-term results: general acquisition of vocabulary.

However, a more detailed analysis taking proficiency and film type into consideration highlights a slightly different picture. Beginner participants took the great-

est advantage from captions, especially in learning the vocabulary in *Harry Potter*, while subtitles seem to have ‘disturbed’ acquisition, as EG2 beginner students fared worse than CG ones (Figure 15), a result that is in line with the reported research on subtitling (Holobow et al. 1984; Danan 1992).

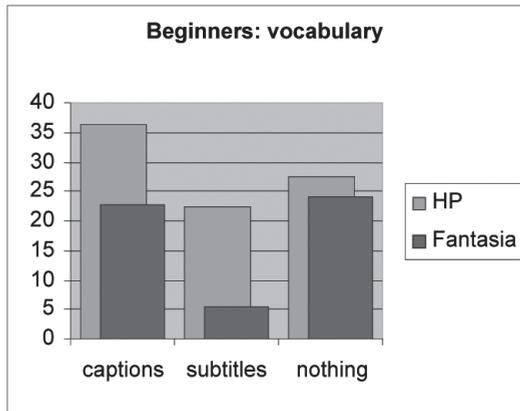


Figure 15. Long-term results: beginners’ results in vocabulary acquisition.

For intermediate students, subtitles are no longer a problem and EG2 results are slightly higher than those of EG1 (Figure 16).

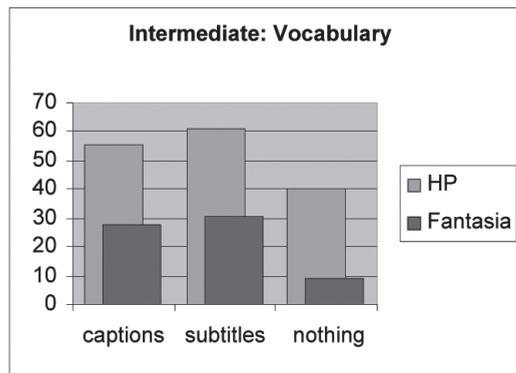


Figure 16. Long-term results: vocabulary results of intermediate students.

Finally, advanced students seem to have taken the greatest advantage from subtitles; as in the case of beginners, the difference between EG1 and EG2 results is more evident with *Harry Potter* clips (Figure 17).

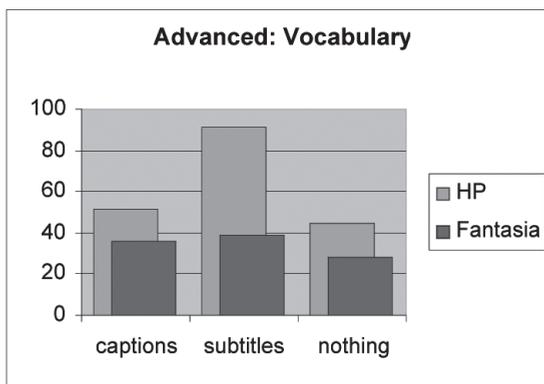


Figure 17. Long-term results: vocabulary results of intermediate students.

Interestingly, all students (regardless of proficiency level or text aid) acquired a greater number of words belonging to *Harry Potter* than to *Fantasia* dialogues (Figure 18), in line with the trend observed in the case of short-term vocabulary comprehension, and the number of words acquired grew with proficiency.

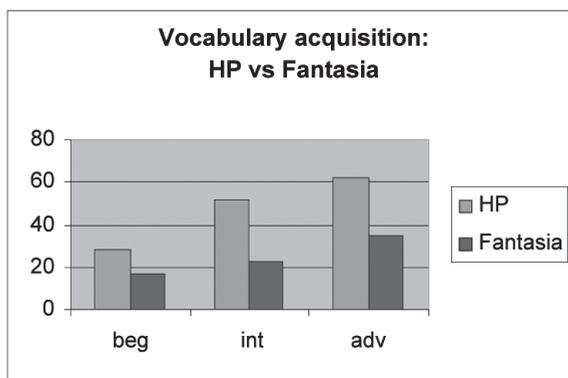


Figure 18. Long-term vocabulary results per film.

3.2.2 Language-in-use

With regard to Task Four (Figure 19), deltas were generally very low, rarely reaching a 25% increase (and this task was composed of only 14 items). In the case of beginner students, text aids did not favour acquisition, as both experimental groups fared much worse than the control group. This result reflects the beginners' trend in language-in-use immediate comprehension. Intermediate and advanced students showed similar trends, with EG2 scoring higher than EG1, which in turn scored higher than CG, a trend that is different from the one highlighted in the case of language-in-use immediate comprehension for these sub-groups.

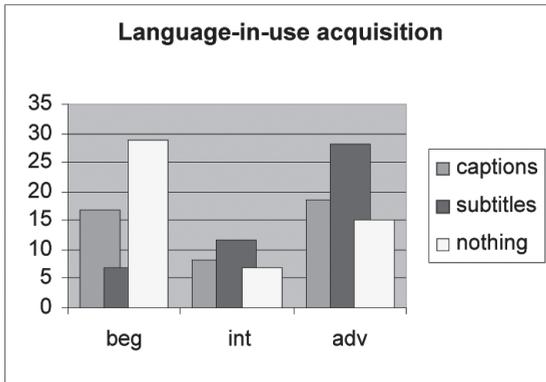


Figure 19. Long-term results in language-in-use acquisition.

A comparison between beginner, intermediate, and advanced students deltas in the two types of film regardless of the presence of textual aids showed the following results (Figure 20): beginners and intermediate students obtained generally higher results with *Harry Potter*, while advanced participants obtained higher scores with *Fantasia*. Unexpectedly, while beginner and advanced students showed consistency with language-in-use immediate comprehension results, intermediate students did not.

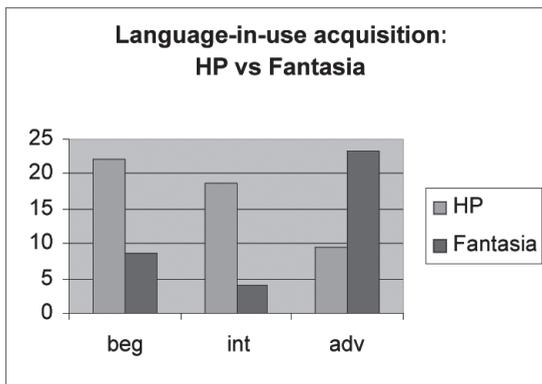


Figure 20. Long-term language-in-use results by proficiency.

4. CONCLUSION

When students watch a film in a foreign language and text aids are displayed, three channels compete in catching the students' attention and in favouring (or hampering) comprehension and learning: one auditory channel, and two visual channels (one verbal and one non-verbal). In this scenario, several different variables are at play, including the following: semantic match between the verbal channels (audio and text) and the non-verbal channel (images); type of text aid (captions, subtitles, no text aid); student level of proficiency; and type of task (content, vocabulary, or language-in-use comprehension or acquisition).

In the current experiment, greater semantic match between audio-video-text inputs helped achieve higher results at all levels of proficiency in short-term comprehension tasks and in both short- and long-term vocabulary tasks, a result that is perfectly in line with previous literature (Baltova, 1994; Duquette & Painchaud, 1996; Grimes, 1990). Comprehension and acquisition of language-in-use, on the other hand, did not consistently benefit from semantic match, especially with higher level students.

As far as text aids are concerned, differences were noticed with respect to type of task and proficiency level. Content comprehension was facilitated by subtitles, immediately followed by captions for beginner and advanced students and by the control situation for intermediate students. In vocabulary comprehension, subtitles proved more useful than captions, especially when proficiency was lower or little or no semantic match existed between verbal and non-verbal channels. The trend was reversed in long-term results, where beginners benefited most from captions, immediately followed by the control situation, while intermediate and advanced students obtained better results with subtitles, immediately followed by captions. Finally, language-in-use comprehension was characterised by a gradual passage from text aids in general and captions in particular limiting comprehension in lower proficiency groups to the complete opposite with advanced students; analogously, language-in-use acquisition was not favoured by text aids when proficiency in English was not very high, but text aids in general and subtitles in particular gradually acquired greater relevance when the proficiency level rose.

In terms of proficiency level, the same proficiency group showed different profiles with respect to the different types of tasks (content comprehension, vocabulary comprehension and memorisation, language-in-use comprehension and memorisation). This may be connected to the intrinsic differences between said activities in terms of nature and cognitive effort. Furthermore, the three proficiency groups benefited to different extents from the various types of text aids: on the whole, beginners were advantaged to a greater degree by subtitles, while more advanced levels gained more advantage from captions. This may partly be due to the fact that subtitles are processed automatically, while captions require a higher level of knowledge of the language before they can be processed without interfering (at least to a minimal extent) with other cognitive processes (listening and taking stock of the video content).

Finally, the different nature of each type of task was made evident by the different profiles across and among proficiency groups. In particular, marked differences emerged between short and long term results for the same type of task. This is probably a consequence of the fact that different processes are involved in short-term and long-term memorisation.

To conclude, the current experiment partially supports the findings described in the relevant literature. A few discrepancies emerged with some previous studies, Danan's (1992) in particular, but they are probably explained by the different type of material and testing procedure adopted. Comparison with previous studies was only possible for short-term content and vocabulary comprehension, and long-term vocabulary acquisition. The language-in-use category was tentatively introduced in this experiment to shift attention towards other important

linguistic issues that had so far been neglected in the literature on subtitling/captioning. However, given the small number of items and the mixed nature of the exercises about the pragmatic use of lexico-grammatical phrases, the results obtained in this category cannot be considered in any way final and further research is needed in this direction.

* University of Salento.

** University of Pavia.

¹ The scene where an off-screen voice introduces the soundtrack (pictured as a single vertical string in the middle of the screen) and

presents the different musical instruments.

² The scenes where the rules of quidditch are explained to Harry and he plays his first quidditch game.

REFERENCES

- Baltova I. (1994), *Impact of video on the comprehension skills of core French students*, in: "Canadian Modern Language Review", 50 (3), 507-531.
- Baltova I. (1999), *Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French*, in: "Canadian Modern Language Review", 56, 32-48.
- Canning-Wilson C. (2000), *Practical Aspects of Using Video in the Foreign Language Classroom*, in: "The Internet TESL Journal", 6 (11).
- Chung J. (1999), *The effects of using video texts supported with advance organizers and captions on Chinese college students' listening comprehension: An empirical study*, in: "Foreign Language Annals" 32 (3), 295-308.
- Danan M. (1992), *Reversed subtitling and dual coding theory: New directions for foreign language instruction*, in: "Language Learning", 42 (4), 497-527.
- Danan M. (2004), *Captioning and Subtitling: Undervalued Language Learning Strategies*, "Meta", XLIX, 1, 67-77.
- Duquette, L., Painchaud G. (1996), *A comparison of vocabulary acquisition in audio and video contexts*, in: "The Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes", 53 (1), 143-171.
- Garza T.J. (1991), *Evaluating the use of captioned video materials in advanced foreign language learning*, in: "Foreign Language Annals", 24 (3), 239-250.
- Grimes T. (1990), *Audio-video correspondence and its role in attention and memory*, in: "Educational Technology Research and Development", 38, 15-25.
- Guillory H.G. (1998), *The effects of keyword captions to authentic French video on learner comprehension*, in: "Calico Journal", 15 (1/3), 89-108.
- Holobow N.E., Lambert W.E., & Sayegh L. (1984), *Pairing script and dialogue: combinations that show promise for second or foreign language learning*, in: "Language Learning", 34 (4), 59-74.
- Koolstra, C.M. & J. W.J. Beentjes (1999), *Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home*, in: "Educational Technology Research & Development", 47 (1), 51-60.
- Lambert W.E., Boehler I., & Sidoti J. (1981), *Choosing the language of subtitles and spoken dialogues for media presentations: Implications for second language education*, in: "Applied Psycholinguistics", 2, 133-148.
- Markham P. (1989), *The effects of captioned television videotapes on the listening comprehension of beginning, intermediate, and advanced ESL students*, in: "Educational Technology", 29 (10), 38-41.
- Neuman S. & Koskinen P. (1992), *Captioned Television as comprehensible Unit: Effects of incidental Word Learning from context for Language Minority Students*, in: "Reading Research Quarterly", 27 (1).
- Paivio A. (1986), *Imagery and Verbal Processes*, New York, Holt, Rinehart & Winston.
- Palladino P. & Bianchi F. (forthcoming). "Improving Foreign Language

Learning at Undergraduate Level: Native language predictive variables and foreign language sensitive skills.”

Price, K. (1983), *Closed-captioned TV: An untapped resource*, in: “MATESOL Newsletter”, 12, 1-8.

Thurstone L.L. & Thurstone T.G. (1981), *PMA: abilità mentali primarie: manuale di istruzioni K-1 (scuola materna e 1. elementare); 2-4 (1., 2., 3. elementare); 4-6 (3., 4., 5. elementare e 1. media); Livello intermedio (11-17) (Batteria fattoriale delle abilità mentali primarie)*, Firenze, Organizzazioni Speciali. [Translation of: Thurstone L.L. & Thurstone T.G. (1963), *Primary mental abilities*].

Vanderplank R. (1988), *The value of teletext subtitles in language learning*, in: “ELT Journal”, 42 (4), 272-281.

Vanderplank R. (1990), *Paying attention to the words: practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) sub-titles*, in: “System”, 18 (2), 221-234.

Vanderplank R. (1993), *A very verbal medium: Language learning through closed captions*, in: “TESOL Journal”, 3 (1), 10-14.