La didattica delle lingue nel nuovo millennio

Le sfide dell'internazionalizzazione a cura di Carmel Mary Coonan, Ada Bier ed Elena Ballarin

Multimodal Input for Italian Beginner Learners of English

A Study on Comprehension and Vocabulary Learning from Undubbed TV Series

Valeria Galimberti (Universitat de Barcelona, Catalunya)

Imma Miralpeix (Universitat de Barcelona, Catalunya)

Abstract Most studies on the use of subtitled videos for EFL learning have focused on intermediate and advanced adult learners viewing short L2 clips with captions. However, there is hardly any that deals with the use of TV series for young learners, and research has rarely assessed the various aspects of vocabulary knowledge in which students may improve when watching TV. In the present experiment, three groups of sixth-grade beginner EFL learners watched a full-length TV series episode under different input modalities: L1 subtitles (N=19), L2 subtitles (N=16), or no subtitles (N=17). All groups showed progress in form and meaning recall from pre-test to post-test, and L2 subtitling led to significantly greater gains in vocabulary recall than L1 subtitling, although this advantage was not shown in the case of episode comprehension.

Summary 1 Introduction. – 2 Literature Review. – 3 Research Questions. – 4 Methodology. – 4.1 Participants. – 4.2 Instruments. – 4.3 Procedure. – 4.4 Data Analysis. – 5 Results. – 5.1 Comprehension. – 5.2 Vocabulary Recall. – 5.3 Meaning Recognition. – 6 Discussion. – 7 Conclusion and Further Research.

Keywords Subtitles. Authentic video. Young learners. Incidental learning. English-as-a-Foreign-Language.

1 Introduction

In the last fifty years, advances in technology have made available to the public a wide range of L2 audiovisual materials and other sources of authentic input for foreign language (FL) learning. At the same time, the function of language in FL classes has gradually shifted from object of analysis to means of communication (Baltova 1999). Accordingly, the use of L2 video to teach foreign languages has been increasingly researched

in the field of second language acquisition (SLA). Some of the key issues under study are viewers' proficiency and age, suitability in the language learning context, and subtitling options (L1 or L2). On the one hand, solid evidence has shown that captions help access authentic FL audiovisuals (Vanderplank 2010; Danan 2004), where the input is mostly comprehensible, i.e. slightly above the viewer's proficiency level (Krashen 1987). On the other hand, it has also been shown that being exposed to contextually rich input without previous or later language activities does not always produce measurable progress (Baltova 1999). In fact, the leisure orientation of TV programs seems to lead to greater engagement and motivation, but it might divert learning efforts from the activity (Vanderplank 2015). The aim of the present study was to collect information about the effectiveness of different input modalities for implicit learning with a sample of 52 low proficiency learners. A TV series episode was used in the attempt to test participants on the type of input they would spontaneously choose in their leisure time that can also promote language learning (Rodgers 2013).

2 Literature Review

The study of subtitled audiovisual materials in relation to various aspects of FL learning is grounded on well-established cognitive theories, such as Paivio's Dual Coding Theory (1986) and Mayer's Theory of Multimedia Learning (2005). A major implication of the Dual Coding approach to SLA is that acquiring L2 vocabulary in association with its nonverbal referents helps retrieving words more efficiently necessary and using them meaningfully (Paivio 1986). Accordingly, one of the main tenets of the Multimedia Learning Theory is that "people learn more deeply from words and pictures than from words alone" (Mayer 2005, 31). However, due the limited nature of working memory, input must be carefully tailored on the students' proficiency, in order to challenge their listening comprehension skills and encourage the activation of both aural and visual processing mechanisms (Sweller 2005).

When these preliminary conditions are satisfied, television programs and TV series in particular can constitute a good source of L2 input in FL learning contexts for a number of reasons. To begin with, they are accessible in large quantities, inviting learners to choose material they are interested in and familiar with (Rodgers 2013). They also tend to create affective engagement, encouraging the viewer to watch many related episodes (Vanderplank 2010). Watching successive episodes, in particular,

 $[{]f 1}$ The terms captions and L2 subtitles (i.e. target language subtitles) are used interchangeably in SLA literature and in this paper.

builds on the same background knowledge of the plot and characters, and also entails lower vocabulary load and higher potential for incidental vocabulary learning than watching unrelated programs (Rodgers, Webb 2011). Moreover, the dialogues of TV series tend to resemble spontaneous conversation more than the speech of other audiovisual genres such as documentaries and news broadcasts. For this reason, comprehension is promoted by the high predictability of interactions that are often based on catchphrases and formulae, and by the high contextualisation of speech within the visual context (Ghia 2012). Finally, there is evidence that the ability to segment speech acquired by watching programs with subtitles can be transferred to other domains of L2 use (Charles, Trenkic 2015).

Regarding the effects of different types of subtitles, a number of studies have proved the learning potential of captioning (Baltova 1999; Mitterer, McOueen 2009; Rodgers 2013), but there is no conclusive evidence that L2 subtitles always support listening comprehension and vocabulary acquisition more than L1 subtitles, or even no subtitles (Vanderplank 2010). According to Mitterer and McQueen (2009), the use of native-language subtitles can be detrimental for learning in L2 speech perception, but other researchers have gathered evidence that it does not impair simultaneous processing of the soundtrack (Danan 2004). In Vulchanova et al. (2015), younger and less experienced learners had better comprehension results with target language rather than native language subtitles, while older and more advanced participants benefited equally from both. It might be the case, then, that different input modalities serve different purposes, also according to the relative stage of second language acquisition learners are at or the skills they want to improve. For instance, Markham, Peter and Mc-Carthy (2001) recommend that lower-level students should watch videos with L1 subtitles first, then with captions and, later on, without subtitles. Sydorenko (2010) points out that watching videos without captions might improve listening comprehension, while captioned videos might be more appropriate for vocabulary learning.

The rationale behind using subtitled TV programs for language learning is that, in order to acquire the necessary vocabulary knowledge to understand authentic FL input, vocabulary learning should not only happen explicitly in the FL classroom, but also incidentally through a substantial amount of exposure to the language (Bisson et al. 2014; Rodgers 2013). In experimental studies on incidental vocabulary learning with authentic L2 audiovisual material, the target words (TWs) are often low frequency words with a high learning burden (Montero Pérez et al. 2014) to reduce the likelihood of the participants being familiar with them. As a consequence, the aural word form might remain unintelligible to students despite a sufficient number of repetitions in the script, especially for less experienced learners. Similarly, meanings might not be easily inferred from the visual context either. In the present study, this problem was ad-

dressed by choosing TWs from all frequency bands on the basis of their saliency in the dialogues. A pre-test was administered to make sure that participants did not know them – and to control for this knowledge if it was the case. In relation to vocabulary learning from videos, Montero Pérez, Van Den Noortgate and Desmet conclude their 2013 meta-analysis stating that more studies are needed on vocabulary learning through captioning and that further research should include different types of vocabulary knowledge tests. According to Nation (2001), assessing the same words with tests at a different level of difficulty, such as recall and recognition, provides richer information on lexical performance. In line with these recommendations, the present study tests three types of vocabulary knowledge: form recall, meaning recall and meaning recognition.

Finally, although the effects of subtitles have been increasingly researched in recent years, short clips have mostly been used, with the exception of a few studies on full-length TV programs (Montero Pérez et al. 2014). Further, the participants in these studies are usually university students at intermediate/advanced proficiency level (e.g. Rodgers 2013; Sydorenko 2010; Yuksel, Tanriverdi 2009), which raises the issue of whether the same effects of subtitling apply to low proficiency learners. A study on listening comprehension by Başaran and Durmuşoğlu Köse (2013) did feature younger participants in grade 8 (age 14). However, the initial proficiency of the participants using L1 and L2 subtitles was lower than that of the no-subtitles group and this, according to the authors, was the reason why the difference between the post-test results of the three groups failed to reach statistical significance.

3 Research Questions

On the basis of the research presented above and considering those issues that remained unexplored, the present study aims at answering the following questions in relation to Italian beginner learners of English:

- 1. Does watching authentic L2 multimodal material with L1, L2 or no subtitles lead to different results in terms of comprehension?
- 2. Does watching authentic L2 multimodal material with L1, L2 or no subtitles lead to different results in terms of vocabulary recognition and recall?

4 Methodology

This study was carried out in a lower-secondary school during the curricular hours of English as a foreign language. Three comparable groups (N=52) of beginner learners watched a TV series episode with either

Italian, English or no subtitles and were tested for comprehension and vocabulary learning. Statistical analysis were performed to see whether vocabulary gains were significant from pre- to post-test, and to assess possible differences between groups.

4.1 Participants

Participants are 52 12-year-old students (30 girls and 22 boys) divided into three groups according to the type of subtitles with which they watched the series' episode (L1S: N=19, L2S: N=16, NOS: N=17). They are at the second year of Italian lower-secondary school (US grade 6) in the same public institute in Milan (Italy), where they take three hours of English classes per week. A Kruskal-Wallis test conducted on the results of an English cloze test did not find significant differences between these groups in terms of proficiency level.

4.2 Instruments

The instruments were adapted from a larger project carried out by the GRAL Group at the University of Barcelona on the use of subtitled series in the classroom. The present experiment included: one 22-minute episode from the first season of the TV series *The Suite Life of Zack and Cody*; a proficiency cloze test; a comprehension test; a vocabulary pre- and posttest based on ten TWs from the video and a multiple choice vocabulary recognition test.

4.2.1 Cloze Test

A cloze test was administered in English to assess L2 proficiency. This test had been used with low-proficiency EFL students in previous studies, showing good validity and reliability results (Muñoz 2006).

4.2.2 Comprehension Test

The comprehension test - written in the participants' L1 - assessed the understanding of the episode through three exercises of different format,

2 In this paper, L1S will be used to indicate the group who watched the episode with Italian subtitles, L2S with English subtitles, NOS without subtitles.

as in Rodgers (2013): (a) true or false, (b) multiple choice, and (c) event ordering. Two versions were administered in each of the three groups, with the same items but in different order.

4.2.3 Vocabulary Pre- and Post-Test (Form and Meaning Recall)

Vocabulary recall was assessed through a pre- and post-test: an aural pretest assessed participants' previous knowledge of ten selected TWs from the episode, presented together with six distractors. Students listened to a native L2 speaker say each word twice, with an interval of three seconds between them, then they were given eight seconds to write down the word, before the voice continued with the following one. At the end, they wrote down the L1 meaning of the words they knew. The post-test was the same as the pre-test, except that it did not include the distractors.

Care was taken to find a balance between frequency of occurrence and linguistic features of the target words included, so that the sample could be representative of the vocabulary found in real language use.

4.2.4 Vocabulary Multiple Choice Test (Meaning Recognition)

For each of the ten TWs in their L2 form, five possible meanings were presented randomly in L1, plus the option: 'I don't know'. The alternatives were: the correct answer (in the same form as it appeared in the L1 subtitles of the TV series episode); a word that belonged to the same frequency band but from a different word class; a word that sounded similar to the target word in L2, but had a different meaning in L1; a word that was similar in L1 meaning and belonged to the same semantic field as the target word; and a hapax legomenon that appeared only once in the episode. Two versions of the multiple choice were devised, with the same content but in different, randomised order. This test was administered after viewing the episode and after the recall test to avoid priming effects in meaning recall.

4.3 Procedure

The experiment took place during the curricular hours of English language classes (two one-hour sessions per group over a period of one week). With each group, the same procedure was followed: in the first session, students completed the cloze test and the vocabulary pre-test. The second session consisted of the treatment (watching the episode under different conditions) and the post-tests.

4.4 Data Analysis

In the comprehension test, one point was given for each correct answer, and partial scores were calculated for each of the three exercises in the test (exercise 1, maximum score: 5; exercise 2, max.: 5; exercise 3, max.: 6). The final scores were a sum of the partial results (maximum score: 16).

In the two aural vocabulary tests (pre- and post-), 0.5 point was awarded for each correct word form and for each correct word meaning (maximum score of 10 in each of the two tests). Then, the pre-test score was subtracted from the post-test score, and the resulting variable was labelled 'vocabulary gains'. The same procedure was followed for word forms only, but each correctly written word form was awarded one point to obtain a comparable maximum score of 10. Then, the pre-test scores were subtracted from the post-test scores to compute L2 'word form' gains. Lastly, a variable named 'word meaning' was computed following the same procedure taking into account the amount of correct L1 word translations.

After obtaining descriptive statistics, Wilcoxon signed-rank tests were performed to see whether vocabulary gains were significant from pre- to post-test. Differences between groups in vocabulary and comprehension were assessed by means of non-parametrical Analysis of Variance tests (Kruskal-Wallis). In case this test showed significant differences, multiple pairwise comparisons were also checked.

5 Results

This section presents the results of the statistical analyses performed on comprehension and vocabulary tests scores. In the vocabulary results, a difference is made between vocabulary recall (form and meaning) and meaning recognition.

5.1 Comprehension

The descriptive data for comprehension seemed to point in the direction of L1S getting the best overall scores, with a mean of 9.89 out of 16, compared to a mean of 8.69 for L2S and 8.82 for NOS. However, Kruskal-Wallis tests on ranks showed no significant differences between the scores of the three groups (p=.264).

5.2 Vocabulary Recall

As described in the next sub-sections, each input modality – L1, L2 or no subtitles – produced small but significant vocabulary recall gains from preto post-test. Participants who had L2 subtitles available achieved higher overall vocabulary recall gains than those having L1 subtitles.

5.2.1 Pre- and Post-Tests

The results of the Wilcoxon signed-rank tests revealed that there were always significant differences between pre- and post-test scores in all groups (see tab. 1).

Table 1. Vocabulary results of the related-samples Wilcoxon signed-rank tests

	Z (p) L1S	Z (p) L2S	Z (p) NOS
Pre vs. post vocab. recall	2.611 (.009)	3.307 (.001)	3.405 (.001)
Pre vs. post word forms	2.490 (.013)	3.275 (.001)	2.996 (.003)
Pre vs. post word meanings	1.998 (.046)	2.844 (.004)	3.231 (.001)

5.2.2 Vocabulary Recall Gains

To specifically control for initial knowledge of TWs, gains were computed for each measure and are presented in table 2.

Table 2. Average vocabulary gains per group

	Mean L1S (SD)	Mean L2S (SD)	Mean NOS (SD)
Vocabulary recall gains	.66 (.944)	1.41 (.953)	1.18 (.809)
Word forms gains	.84 (1.302)	1.44 (.892)	1.35 (1.222)
Word meanings gains	.47 (.964)	1.38 (1.310)	1.00 (.866)

The Kruskal-Wallis test yielded significant results for overall vocabulary recall gains ($H_{(2)}$ =6.467, p=.039). Multiple pairwise comparisons showed that L2S got significantly higher scores (p=.018) than L1S, but not than NOS. It should be noted that the p value of the comparison between L1S and NOS was close to significance (p=.055). The separate non-parametric Analysis of Variance tests for word form recall gains and meaning recall gains gave non-significant results.

5.3 Meaning Recognition

Descriptive statistics indicate that L1S got the highest average score in the meaning recognition post-test: 6.47 out of 10 TWs, compared to 6 TWs (L2S) and 5.59 TWs (NOS). However, the Kruskal-Wallis test found no significant differences between the three groups (p=.158).

6 Discussion

In this experiment, each of the three conditions lead to small but significant vocabulary gains, in spite of the low proficiency of the participants, the shortness of the treatment and the incidental learning condition. The L2 subtitles group significantly outperformed the L1 subtitles group on overall vocabulary recall, but the difference was not significant when form and meaning recall gains were considered separately. There was no significant between-group difference for comprehension scores and meaning recognition either.

Regarding comprehension, the findings of the present study do not fall in line with the results of Vulchanova et al. (2015), Markham, Peter and McCarthy (2001) and Baltova (1999), where the groups watching the video with captions or subtitles significantly outperformed the 'no captions' group. However, this might be due to the very low proficiency and young age of the participants in the present experiment: they may not be ready to make the most of the subtitles to gain better comprehension, or they may not be able to read them all fast enough, as students are not usually familiar with this practice. As mentioned above, Başaran and Durmuşoğlu Köse's study (2013) also used a 20-minute video with either L1, L2 or no subtitles and failed to find significant differences in comprehension between conditions, a result that was attributed to differences in proficiencies between the groups. Although this limitation was addressed in the present study by choosing participants with the same EFL instructional background and controlling for between-group differences with a proficiency test, no difference was found between conditions in this case either.

The finding that L1 subtitles did not promote short term L2 vocabulary learning more than the other two conditions is not in line with some previous experimental studies that found the use of L1 subtitles to be beneficial in terms of L2 soundtrack processing (Danan 2004; D'Ydewaelle, Van de Poel 1999). However, the results of the present study would agree with the claim that L1 subtitles may hinder the association of phonological information in the L2 with the relative written forms and, therefore, might interfere with perceptual learning in speech (Mitterer, McQueen 2009). Another possible explanation is that, as in the study by D'Ydewaelle and Van de Poel (1999), our young students learned more from soundtrack than from subtitles, whereas the opposite usually happens with adult learners. Since the participants in

the study by D'Ydewaelle and Van de Poel were between 8 and 12 years old, their conclusions can be applicable to the present study too. Giving the very few available studies with young beginner learners and multimodal materials, this is a question that definitely requires further investigation.

In spite of the low means, there were significant recall gains for all three groups. Similarly, significant gains – but no significant difference between captions and no captions conditions – were found by Karakas and Saricoban (2012), and Yuksel and Tanriverdi (2009), two studies that used the VKS³ as a pre- and post-test. In an attempt to deal with the limitations of self-report testing instruments, the present experiment included a variety of tasks that objectively tapped into several types of vocabulary knowledge. However, the between-groups comparison of all measures of vocabulary acquisition tested showed marginal between-group differences only.

7 Conclusion and Further Research

In this study on the effects of viewing authentic tv series with different types of subtitles, each input modality -L1, L2 or no subtitles- led to small but significant vocabulary gains, which gives evidence that watching this type of material can enhance language learning also at beginner levels. L2 on-screen text was found to be significantly more beneficial than L1 subtitling in terms of vocabulary recall, which supports previous studies on the effectiveness of L2 subtitling for speech segmentation and phonological mapping (Mitterer, McQueen 2009; Charles, Trenkic 2015). Nevertheless, the role of authentic audiovisual materials for language learning in young beginner learners needs to be extensively explored.

Further research on the use of authentic L2 multimodal input is also needed to investigate the possible long-term effects of this type of classroom intervention for implicit language learning at different proficiency levels. Future studies should analyze with longitudinal designs the different aspects of receptive and productive vocabulary knowledge that can be enhanced with this practice.

Acknowledgements

We would like to thank the professors, headmaster and students of the school I.C. Settimo Milanese for their participation and support, and project FFI2013-47616-P from the Spanish Science and Innovation Ministry to the GRAL Research Group.

³ The Vocabulary Knowledge Scale (VKS) is a self-report form of assessment that requires the student to provide linguistic information about the words s/he knows, such as translations or examples of word use in context (Paribakht, Wesche 1996).

References

- Baltova, Iva (1999). The Effect of Subtitled and Staged Video Input on the Learning and Retention of Content and Vocabulary in a Second Language [Doctoral Thesis]. Toronto: University of Toronto. URL http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp04/nq41096.pdf (2018-01-18).
- Başaran, Hülya Felek; Durmuşoğlu Köse, Gül (2013). "The Effects of Captioning on EFL Learners' Listening Comprehension". *Procedia, Social and Behavioral Sciences*, 70, 702-8.
- Bisson, Marie-Josée; Van Heuven, Walter J.B.; Conklin, Kathy; Tunney, Richard J. (2014). "The Role of Repeated Exposure to Multimodal Input in Incidental Acquisition of Foreign Language Vocabulary". *Language Learning*, 64(4), 855-77.
- Charles, Tendai; Trenkic, Daniela (2015). "The Effect of Bi-Modal Input Presentation on Second Language Listening: The Focus on Speech Segmentation". Gambier, Yves; Caimi, Annamaria; Mariotti, Cristina (eds.), Subtitles and language learning. Bern: Peter Lang, 173-97.
- Danan, Martine (2004). "Captioning and Subtitling: Undervalued Language Learning Strategies". *Meta*, 49(1), 67-77.
- D'Ydewalle, Géry; Van de Poel, Marijke (1999). "Incidental Foreign-Language Acquisition by Children Watching Subtitled Television Programs". *Journal of Psycholinguistic Research*, 28(3), 227-244.
- Ghia, Elisa (2012). Subtitling Matters. New Perspectives on Subtitling and Foreign Language Learning. Oxford: Peter Lang.
- Karakas, Ali; Saricoban, Arif (2012). "The Impact of Watching Subtitled Animated Cartoons on Incidental Vocabulary Learning of ELT Students". *Teaching English with Technology*, 12(4), 3-15.
- Krashen, Stephen D. (1987). *Principles and Practice in Second Language Acquisition*. New York: Prentice-Hall.
- Markham, Paul. L.; Peter, Lizette A.; McCarthy, Teresa J. (2001). "The Effects of Native Language vs. Target Language Captions on Foreign Language Students' DVD Video Comprehension". Foreign Language Annals, 34. 439-45.
- Mayer, Richard E. (2005). "Cognitive Theory of Multimedia Learning". Mayer, Richard E. (ed.), *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press, 31-48.
- Mitterer, Holger; McQueen James M. (2009). "Foreign Subtitles Help but Native-Language Subtitles Harm Foreign Speech Perception" [online]. *PLoS ONE*, 4(11), e7785. DOI 10.1371/journal.pone.0007785.
- Montero Pérez, Maribel; Van Den Noortgate, Wim; Desmet, Piet (2013). "Captioned Video for L2 Listening and Vocabulary Learning: A Meta-Analysis". *System*, 41(3), 720-39.

- Montero Pérez, Maribel; Peters, Elke; Clarebout, Geraldine; Desmet, Piet (2014). "Effects of Captioning on Video Comprehension and Incidental Vocabulary Learning". *Language Learning and Technology*, 18(1), 118-41.
- Muñoz, Carmen (ed.) (2006). *Age and the Rate of Foreign Language Learning*. Clevedon: Multilingual Matters.
- Nation, Paul (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Paivio, Allan (1986). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.
- Paribakht, Tahereh Sima; Wesche, Marjorie Bingham (1996). "Enhancing Vocabulary Acquisition through Reading: A Hierarchy of Text-Related Exercise Types". *The Canadian Modern Language Review*, 52(2), 155-78.
- Rodgers, Michael (2013). English Language Learning through Viewing Television: An Investigation of Comprehension, Incidental Vocabulary Acquisition, Lexical Coverage, Attitudes, and Captions [Doctoral Thesis]. Wellington: Victoria University of Wellington. URL http://researcharchive.vuw.ac.nz/handle/10063/2870 (2018-01-18).
- Rodgers, Michael; Webb, Stuart (2011). "Narrow Viewing: the Vocabulary in Related Television Programs". TESOL Quarterly, 45(4), 689-717.
- Sweller, John (2005). "Implications of Cognitive Load Theory for Multimedia Learning". Mayer, Richard E. (ed.), *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press, 117-33.
- Sydorenko, Tetyana (2010). "Modality of Input and Vocabulary Acquisition". Language Teaching and Technology, 14(2), 50-73.
- Vanderplank, Robert (2010). "Déjà Vu? A Decade of Research on Language Laboratories, Television and Video in Language Learning". *Language Teaching*, 43(1), 1-37.
- Vanderplank, Robert (2015). "Thirty Years of Research into Captions/ Same Language Subtitles and Second/Foreign Language Learning: Distinguishing between "Effects of" Subtitles and "Effects with" Subtitles for Future Research". Gambier, Yves; Caimi, Annamaria; Mariotti, Cristina (eds.), Subtitles and Language Learning. Bern: Peter Lang, 19-40.
- Vulchanova, Mila; Eshuis, Hendrik; Aurstad, Lisa M.; Kvitnes, Ingrid E. (2015). "As Naturalistic as It Gets: Subtitles in the English Classroom in Norway". *Frontiers in Psychology*, 5, 1-10.
- Yuksel, Dogan; Tanriverdi, Belgin (2009). "Effects of Watching Captioned Movie Clip on Vocabulary Development of EFL Learners". *The Turkish Online Journal of Educational Technology*, 8(2), 48-54.