



LEARNING ABOUT SCHOOLS IN THE BRITISH ISLES THROUGH A VIDEO CORPUS: REFLECTIONS ON AN ONLINE PROJECT FOR DIGITAL LITERACY AND MULTIMODAL CORPUS CONSTRUCTION

MARIAVITA CAMBRIA

Potrjeno/Accepted

20. 6. 2023

University of Messina, Department of Ancient and Modern Civilizations, Messina, Italy

Objavljeno/Published

17.7. 2023

CORRESPONDING AUTHOR/KORESPONDENČNI AVTOR

mcambria@unime.it

Abstract/Izvleček

Through first-hand experiences undertaken with students using the OpenMWS platform (Taibi 2020), the paper reports on the progress made by second-year language degree students in English linguistics at the University of Messina as regards their use of online corpus construction, annotation and search tools when exploring video genres. The paper describes the ways in which participation in the Messina OVP (Online Video Project) on schools in the British Isles during and after the Covid crisis has proved beneficial for the students in question, both in terms of acquiring textual competences and as regards creating new interactive communities (Vasta and Baldry 2020; Baldry et al 2022).

Keywords:

Multimodality, corpus construction and annotation, digital literacy, English linguistics, video genres.

Ključne besede:

večkodnost, gradnja in označevanje korpusa, digitalna pismenost, angleško jezikoslovje, video žanri.

Učenje o šolah na Britanskem otočju z video korpusom: refleksija o spletnem projektu za digitalno opismenjevanje in večkodno gradnjo korpusa

Na podlagi pridobljenih izkušenj s študenti, ki uporabljajo platformo OpenMWS (Taibi 2020), članek poroča o napredku študentov drugega letnika študija angleškega jezikoslovja na Univerzi v Messini pri uporabi spletne gradnje korpusa, označevanja in iskalnih orodij za raziskovanje video zvrsti. Prispevek opisuje načine, ki so se pri sodelovanju projekta Messina OVP (Online Video Project) na šolah na Britanskem otočju med kovidno krizo in po njej za učence izkazali kot koristni, tako v smislu pridobivanja besedilnih kompetenc kot v zvezi z ustvarjanjem novih interaktivnih skupnosti (Vasta in Baldry 2020; Baldry et al 2022).

UDK/UDC:

37.091.3

DOI <https://doi.org/10.18690/rei.16.Sp.Iss.2988>

Besedilo / Text © 2023 Avtor(ji) / The Author(s)

To delo je objavljeno pod licenco Creative Commons CC BY Priznanje avtorstva 4.0 Mednarodna.

Uporabnikom je dovoljeno tako nekomercialno kot tudi komercialno reproduciranje, distribuiranje,

dajanje v najem, javna priobčitev in predelava avtorskega dela, pod pogojem, da navedejo avtorja

izvirnega dela. (<https://creativecommons.org/licenses/by/4.0/>).



University of Maribor Press

Introduction

Using videos for language learning has gained ground in the last few decades, becoming a crucial part of the language learning process (Montero Perez and Rodgers, 2019). Videos can be adapted for educational use for students of different ages, thus raising language awareness and making students more active in their language learning process. The potential of Videos for second language learning is further highlighted by studies on the role of out-of-class exposure for language learning especially in mobile learning. Data on out-of-school activities show that viewing videos in a foreign language is a popular out-of-school activity also for young learners (Vasta 1996; De Wilde, Brysbaert and Eyckmans 2019).

The current research paper was carried out within the Messina OVP (Online Video Project), a two-year project undertaken at the University of Messina with students enrolled in Year II of the course in English Language and Linguistics which is part of the language degree course at the University of Messina. The project began during the first lockdown in March 2020 and was offered to students enrolling in the 19-20 and 20-21 academic years. Of the almost 350 students enrolled in the course over the two years, 172 students decided to take part in the project and 164 completed all the tasks. The OVP project consists of three subcorpora, the result of the students' work, respectively named: 1. OVP_A (2020) which is about "Covid"; 2. OVP_B (2021) about "Covid in schools" and 3. OVP_C (2021) about "The School System in the British Isles". The main goal of OVP is to raise awareness in foreign language learning processes, mainly English as a Foreign language (EFL), thanks to the pedagogical enhancement of video corpus construction and annotation as educational tools. It is also designed to foster learners' textual competences in English as part of their undergraduate course.

The field of educational use of corpora for language learning has grown considerably in the last twenty years (Jablonkai and Csomay 2022) due to the shared efforts of researchers and practitioners who have looked at how the active task of creating, annotating and accessing corpora among foreign language learners has impacted their foreign language and communication skills (McEnery & Wilson 1997; Frankenberg-Garcia, Flowerdew-Garcia et al. 2011; Boulton & Cobb 2017; Poole 2022).

The active involvement of learners accessing corpora for language learning refers to accessing already-made corpora via tools in order to explore, inter alia, vocabulary, grammar, collocations of words and sentence structure (Liu & Jiang 2009; Nation 2001). Several works and projects have paved the way in the field of educational use of corpora (Bernardini 2003) and, as argued by Poole, studies have shown that “not only are learners able to access authentic language use for their various target communities, they are able to do so independently in moments where such discoveries are most timely relevant and rewarding” (2022: 1623).

Within research encouraging access and consultation of corpora by students (Granger 2015; Poole 2018), active corpus construction by them has received little attention (Flowerdew 2015) where this is understood to refer to students’ production of corpora to enhance their lexical, writing, listening comprehension and even oral skills, but also research and collaboration skills via peer work (Baldry and Thibault 2020; Baldry and Kantz 2020; Cambria et al 2012). Once the corpus is produced, this type of use shares similar educational features with large, lemma-based corpora for students (such as BNC and COCA), but there are quite a few beneficial effects associated with the process of corpora construction (Baldry and Thibault 2006, 2020; Ackerley and Coccetta 2007). Indeed, students need to actively choose their texts and this has several beneficial outcomes in their learning process. When creating video corpora, they also have to unpack videos and be able in this process of disassembly, to recognize and untangle the various semiotic resources. Additionally, they take on an active role in revision, potentially in collaboration with their peers. Teachers often note that one of the major problems facing teachers in the use of corpora is the difficulty of finding corpora and language samples suitable for learners (Breyer 2009) and the lack of available and learner-accessible corpora (Frankenberg-Garcia 2012). However, this type of corpus use may be seen as overcoming some of teachers’ complaints about the amount of time required to find and/or build a suitable and appropriate corpus and, hence, the time-consuming nature of corpus study, an issue first addressed by Johns (1991).

The aims of the current study are thus twofold. On the one hand, it seeks to look at the impact that the use of videos has on mediation and language students in their awareness about linguistic components. On the other hand, it aims to investigate whether corpus construction and annotation leads to an improvement in learners’ digital skills and to the promotion of a critical awareness in accessing texts.

This includes overcoming some of the main problems encountered in corpus-aided pedagogy. In this respect, this paper focuses on the last stage – Stage 5 Language analysis of the 2021 OVP_C corpus “The School System in the British Isles” – which included 19 videos and saw the participation of 60 students by answering the following research questions:

(RQ1) To what extent can creating one’s own video corpora improve a student’s language competences?

(RQ2) Do students improve their knowledge of how cohesion and coherence work in English texts via a direct approach to textuality?

(RQ3) To what extent can the educational use of corpus construction and annotation foster students’ textual awareness?

(RQ4) To what extent does the educational use of videos foster multimodal and digital competences?

After an overview of the Messina OVP project given in Section 2 (essential to understand the rationale behind Stage 5 “Language analysis”), in Section 3 the materials and methodology used in Stage 5 are illustrated, while Section 4 discusses the results followed by some conclusions.

The Online Video Project: context and background info

The Messina OVP project is one of a number of research projects using the OpenMWS platform for the construction of video corpora, co-designed by Davide Taibi and Anthony Baldry, and implemented by Davide Taibi at ITD-CNR, Palermo (Taibi 2020; Baldry et al 2020; Baldry and Thibault 2020). One of the basic functions of OpenMWS is to allow any YouTube video to be repurposed as part of online video corpora by students in their very first years of university study.

As mentioned above, the Messina OVP project is a two-year project which began during the first lockdown in March 2020. As part of their course assignment (English Language and Linguistics II year), students were asked to create small study groups (3-5 people in each group) and select YouTube videos on a specific topic (“The Covid Pandemic in the British Isles” for the 2019-2020 student cohort; “Covid in Schools” and “The School System in the British Isles” for the 2020-2021 student cohort).

Folders functioning as pigeon holes were created in a Google Drive for each group's five-stage task of creating an online multimodal transcription of their selected video and an oral presentation thereof. Each stage was presented during a meeting with the teacher, and students were given a precise deadline for each stage which was then followed up by the teacher's feedback and assessment. A total of 8 teacher-group meetings were held for each of the two student cohorts in the overall project which lasted from four to eight months (March-June 2020 and October-May 2021). Each group was asked to choose a group leader so that communication could be quicker. This step-by-step procedure allowed the students to be monitored in their work and prompted group discussions around the problems that arose during each individual stage. It was also a way to create group cohesion and group discussion with the students being assessed both for their individual and group work.

Working entirely online, students followed the step-by-step procedure of the Messina OVP stages and undertook activities that exercised their digital and linguistic skills in relation to three types of online tools: a. Microsoft Teams for real-time teacher-group, student-student and group-group interaction b. Google Drive for sharing and handing in the different stages of the work c. YouTube for access to videos and autogenerated transcriptions. Students were also asked to use word processing and spreadsheet tools during the stages leading to the final oral presentation in which each group presented and discussed its video during a seminar. The project work required each group of students to record metadata for the video they selected in a shared file accessed via a group folder in the project's Google Drive that contained all the information about the video they had selected.

Indeed, the project requires students to master digital skills in keeping with CFRIDiL proposals (Sindoni et al, 2019) so that knowing the basics of software packages are an essential requirement that the project exercises to the full. The manuals created by Kantz (Baldry and Kantz 2022), in the initial stages of the development of OpenMWS were adapted by the author to the needs of the Messina OVP Project.

In the first stage, "Video selection", students were asked to select the video and divide it equally among the group members so that each had approximately 5 minutes' video to handle and analyse. They were given precise instructions to carry out their Stage 1 tasks with its various steps which required them to cope with choosing an appropriate video (average length is 20-30 minutes), understand what some specific technical terms meant (for instance 'YouTube video short identifier') and use them appropriately.

Having selected their videos, in the second stage, the “Transcript” stage, students were asked to download YouTube’s auto-generated transcripts. Once these had been downloaded, they were asked to copy and paste the text into a word processing document, They then had to reorganise the downloaded text so that it corresponded in an intelligible way to the video soundtrack. This restructuring involved intensive listening, viewing and punctuation skills. Through this step they gained a general idea of what restructuring a video into its component micro and macro-units entails (Baldry and Thibault 2006). In terms of comprehension, this step encourages a sense of how and why automatic transcription can be just as wrong as automatic translation (Perego and Taylor 2022).

In Stage 3, the students were introduced to the division of a video into sequences and what a sequence is. In other words, they had to divide the text up into units of meaning making decisions about formal and functional aspects as they did so. They were also asked to produce coherent divisions into meaningful sub-units. In Stage 4, they had to recognize the different semiotic resources used by filling in Table 1 using spreadsheet tools.

Table 1: OVP Stage 4 columns

Group 3	Sequence	Time point	Time span	Oral discourse	Sounds	Written discourse	Visual sequencing
Item 4	Sequence 1	00:00	00:06	None	Music	Title: cincinnati children’s changing the outcome together	Fr1: cincinnati children’s icon with hospital behind

As shown in Table 1, they had at their disposal a column for Oral Discourse, one for Sounds, Written Discourse and Visual Sequencing. For beginners in online transcription this is more than enough, as one of the main aims of the project is to encourage transcription skills in relation to disentangling and correcting an imperfect written text.

Stage 4 also introduces tags and metacharacters that can be used for dedicated search procedures designed to understand how the various semiotic resources interact in the meaning-making process. Specifically, it leads students into recognizing the different semiotic resources such as, for instance, oral vs written discourse and the distribution of their roles in a video. Tables 2 and 3 are examples taken from a video “Back to school in the age of Covid”, part of the OVP_C corpus entitled “The School System in the British Isles”, in which a female speaker is explaining the situation in schools at the time of COVID. In Sequence 2, in the Oral Discourse Column, the tag ONFS stands for ONscreen Female Speaker while the written discourse indicates that there is a caption with the name of the place, i.e., Loreto College Mullingar. The Visual Sequencing column identifies what can be seen in the video, in this case Sharon Tobin standing in the PE hall.

Table 2: Sequence 2 video “Back to school in the age of Covid”

ORAL DISCOURSE	SOUNDS	WRITTEN DISCOURSE	VISUAL SEQUENCING
ONFS, Sharon Tobin: It’s been around 170 days since secondary school students were in the classroom. Since then, they’ve been engaged in distance learning and, of course, they’ve had their summer holidays too, but this week thousands of them have started to return to the classroom and it’s been a very different experience. We met up with some of them to find out how they’re coping with education in the time of COVID.		CAPTION: Loreto College Mullingar	FR1: Sharon Tobin standing in the PE hall

Sequence 3 (see Table 3) shows changes both in the tags used and the overall orchestration of semiotic resources.

Sharon Tobin is off screen (thus the tag indicates OFFS, i.e. OFscreen Female Speaker), there is the ambient sound of the sanitiser being clicked as well as signage, i.e. a piece of written discourse which is clearly visible on doors, equipment or other objects as, for example, the instruction on how to use the hand sanitiser. In this sequence, the visual image is made up of a series of frontal shots with the camera panning from the hand sanitizer to the classroom. Unlike Sequence 2, in Sequence 3 viewers get to know which school the children attend from the speaker's words rather than from an internal diegetic caption.

Table 3: Sequence 3 video "Back to school in the age of Covid"

ORAL DISCOURSE	SOUNDS	WRITTEN DISCOURSE	VISUAL SEQUENCING
OFFS, Sharon Tobin: Sanitising hands throughout the day, masks in the classroom and lockers that are out of bounds. The new reality for students here at St. Joseph secondary school in Rochfortbridge. Like most schools, it's opening on a phased basis. Today the sixth-year students were in.	AMBIENT SOUND: Hand sanitiser click	SIGNAGE: Coronavirus COVID-19 How to use hand sanitiser	FR2: Close-up of a hand sanitiser dispenser; FR3: Students in a classroom; FR4: Close-up of lockers; FR5: Close-up of a teacher in the classroom; FR6a: Students in a classroom; FR6b: Close-up of two students

Methodology

Stage 5, the language analysis stage, which is the main focus of this paper, was tailored for language students and was directly related to the syllabus stating that:

“[...] the course will portray the English language from multiple perspectives providing students with the necessary grounding in linguistics to encourage independent investigation especially in sentence structure. Key concepts such as cohesion and coherence, grammar and grammaticality will be introduced. Students will be brought to the centrally important core texts with a focus on the multisemiotic production of meaning but with a strong emphasis on how texts are built and understood. Particular emphasis will be placed on the various semiotic resources used in the production of “static” (i.e. printed) and “dynamic” (i.e. websites, movies etc.) texts in English.”

Stage 5 of the project focused precisely on developing an awareness about the work carried out in Stages 1-4, and about metalinguistic skills that delved, inter alia, into aspects of cohesion and coherence. The aim was also to learn specialized vocabulary deductively, in the case of the OVP_C corpus that of the School system in the British Isles. Figure 1 shows the details of Stage 5, which was divided into 8 steps.

2020-2021 Messina Online Video Project

**MESSINA ONLINE VIDEO PROJECT
STAGE V**

Stage IV produced an Excel document in horizontal layout with a table containing 11 columns: GROUP, SEQUENCE, TIME POINT, TIME SPAN, ENDPOINT, ORAL DISCOURSE, SOUNDS, WRITTEN DISCOURSE, VISUAL IMAGES, GLOSSARY AND NOTES. Please now use this file to comment on several aspects and features of your video. First give some general information about it and then give your detailed comments. Please remember to rename the file as "OVP_Stage V_Group number"(ex. OVP_STAGE 5_901) and to upload it in your group folder.

TITLE OF VIDEO AND LINK: _____
DURATION: _____
DATE OF PRODUCTION: _____

STEP 1 SHORT SUMMARY OF THE VIDEO

Please provide a short summary of your video (no more than 300 words)

STEP 2: TRANSCRIPT AND SEQUENCES

Please comment on the auto-generated subtitles, your transcript and on the sequences used (including the visual ones) in no more than 150 words.

STEP 3: LEXICAL AND GRAMMAR COHESION

In no more than 300 words, please identify the items used in the video to create cohesion commenting in particular on:

1. Grammar cohesion (personal/possessive pronoun reference, deictic references, comparative structures etc.);
2. Lexical cohesion (repetitions, synonyms, antonyms, superordination, ordered series, specific/general reference, whole/part, meronyms, semantic field, homophones etc.);

STEP 4: LEXICAL ELEMENTS

Please comment on the video's lexical elements (morphology, acronyms, word formation processes, conversion, compounds etc.) in no more than 250 words

STEP 5: THE VERB SYSTEM

Please comment on tense, aspect, modality and voice (no more than 250 words).

STEP 6: ORAL DISCOURSE

Please comment on significant items in the oral discourse including the use of sounds (no more than 250 words).

STEP 7: WRITTEN DISCOURSE

Please comment on significant items in the written discourse (use of captions, titles etc.) in no more than 250 words

STEP 8: OTHER COMMENTS

Please use this space for other comments on your video and on the project (no more than 500 words)

Manual created by Prof. Mariavita Cambria. Based on the OpenMWS manuals design developed by Prof. Deirdre Kantz.

Figure 1: OVP Stage 5 (language analysis)

Step 1, the summary, allowed students to focus on the main topic of the video by isolating specific vocabulary related to the video itself. A new interface named Summary and Peer Review has also been recently added to OpenMWS partly as a result of the Messina OVP project where group work was a fundamental parameter. This allows summaries of each of the 19 videos making up the corpus to be accessed either by browsing sequentially (using the Prev. Item or Next Item navigation resources) or, more directly, by selecting a specific Video Item from the Tree shown in the top left-hand corner of Figure 2. The system also includes the multi-summary section which provides for comparative access to all the summaries and language analysis in the project. In this way, students had access to a shared repertoire containing the lexicon used in the various videos which encouraged them to expand their knowledge about the topic.

The screenshot displays the 'Multi-summary and Peer Review' interface. On the left, a vertical list of 19 items is shown, each with a folder icon and a label like 'ITEM MC_V10 27_10_22_01'. The main area on the right features a navigation menu with options: Summary, GROUP, STEP 2_TRANSCRIPT, STEP 3_COHESION, STEP 4_LEXICAL ELEMENTS, STEP 5_THE VERB SYSTEM, STEP 6_ORAL DISCOURSE, STEP 7_WRITTEN DISCOURSE, and STEP 8_OTHER COMMENTS. Below this is a summary box for 'MB 03 - Back to school in the age of Covid-19' with a text excerpt. At the bottom, a metadata table provides details for the selected item.

UNMESSINA xstm_x2 - ITEM MB_V10 27_10_22_03		
<< Prev Item		
Item:	Duration:	Video:
ITEM MB_V10 27_10_22_03	04:15	mbf0TQV41BE
Video title: Back to school in the age of Covid-19 Duration: 04:15 Excerpt: Target Audience: General public Perspective: School management after COVID-19 Organization: RTE News Type of film: Interview Date: 24.08.2020 Produced By: RTE News Place: IRL Subgroup: IJ_SUBCORPUS B Subgroup Topic: COVID_8_SCHOOLS		

Figure 2: The Multi-Summary and Peer Review in the OpenMWS interface

In the “Transcript and sequences” step, students were asked to actively discuss their decisions made in correcting the transcript which often implied checking the vocabulary and grammar used. They were also asked to comment on the rationale used to divide the video into sequences, a way to prompt active reflection on the concept of video-sequence. Step 3 tested students’ ability to differentiate between grammatical and lexical cohesion. In grammatical cohesion, they were asked to isolate subcategories, such as co-referential pro-forms (e.g., they, this or there), substitution (which covers grammatical placeholders e.g., one, do or so), ellipsis, conjunctions (which can be realised by simple grammatical means e.g., or, and or but or by more complex lexical expressions e.g., as a result). Their lexical competence was thus tested in the recognition and listing of lexical items by means of literal repetition, synonymy or hyponymy and collocation through lexemes that typically co-occur in texts. Step 4 aimed at checking students’ ability to access authentic language use and recognise word-formation processes typical of the English language. This type of competence is particularly relevant for language students as it helps in creating autonomous and independent access to discourse in English of all types.

Step 5 was tailored to help students familiarize with the English verb system and, in particular, with modality which constitutes an important pillar in the English Language and Linguistics Year 2 written exam. Steps 6 and 7 invited reflection on the selected videos’ use of oral and written discourse. Students were invited to reflect on how features of oral discourse (e.g. accent, pronunciation, conversational and prosodic features) and written discourse (e.g. captions, diegetic subtitles etc.) worked together in the meaning-making process. Finally, Step 8 allowed students to comment on the project as a whole by giving feedback on positive and negative outcomes on the use of videos and on corpora construction.

Discussion and conclusion

60 students divided into 19 groups worked specifically on the construction of the OVP_C corpus “The School system in the British Isles”. They all carried out Stage 5 as group work, thus the sample of results from Stage 5 consists of 19 examples compiled in all 8 steps.

A mixed approach was used to investigate the results and answer the research questions: a) quantitative analysis of the results of the students during their English Language and Linguistics II written and oral assignments in the project and in their final exams (written and oral exams); b) qualitative analysis of each group's Stage 5 work.

As for the OVP assignment, the drop-out rate was very low: only 2 students did not sit for the final assignment, while 58 completed it successfully. As for the final exam, it is useful to add that the English Language and Linguistics (II year) exam includes a written and an oral test and that at the University of Messina language students have, at the time of writing this paper, 4 written and 10 oral sessions for their exams during each academic year. The written test is a CEFR B2 exam, the format is made up of: 1. use of English; 2. reading comprehension 3. listening comprehension and 4. a written production. 85% (n = 51) of the students who took part in the project passed the written exam at their first attempt. As for the oral exam, 95% (n = 57) passed their oral exam at their first attempt. The above-mentioned data were collected in March 2022, i.e., when students had the last chance to sit for the exam. With regard to the RQ1 ("To what extent can creating one's own video corpora improve a student's language competences?"), the qualitative analysis of data confirms the exam data, as the students wrote in their comments that they felt that their listening and deductive skills had improved. All the groups emphasized the beneficial effect of working with auto-generated transcripts in terms of an improvement in: a) their use of punctuation; b) their listening abilities. Some of the groups also mentioned British accents and their 'effects' on the auto-generated transcript. The following comments are an example of the attitude towards listening and punctuation:

"[...] we have also improved our listening abilities; in fact, we had to listen to the video numerous times to get the words." (Group 921)

"You always end up learning new elements and you will remember the experience forever. This happens regardless of whether you encounter difficulties or not. Actually, we admit that at the end of the work some of the things were extremely beneficial. We learnt different aspects of English language, from vocabulary to new expression, to a more accurate way of listening." (Group 925)

“We found a modest number of mistakes due mainly to the strict pronunciation and the strong Irish accent of the speaker.” (Group 922)

Some groups also discussed in an extremely conscious way the concept of “voice”, as revealed in the following comment:

“Finally, all the discourse is expressed through active form but there are exceptions in which the passive form is used to focus the attention on children’s work and on the help they receive from teachers, for instance when the narrator says: “the extra help Max is received from Kate.”. (Group 929)

In order to explore RQ2 (Do students improve their knowledge of how cohesion and coherence work in English texts via a direct approach to textuality?), it is important to clarify that students had 20 hours of lectures on these concepts and were asked to read extracts taken from Halliday and Hasan’s *Cohesion in English* (1976), before being asked to analyse them independently. Steps 3 and 4 were carried out very well by all the students with an appropriate use of specific words and a clear understanding of how each category contributes to text construction and meaning-making. The use and function of personal pronouns, deictics, the basics of English morphology, an awareness of main word formation processes and of the role of acronyms in the English language emerge from following comments:

“Deictics is another frequent device used to create cohesion, they are especially time deictics. “Yesterday” and “last week” make reference to past events, while “today” and “now” are used to relate to present events.” (Group 925)

“There are also some technical jargon of university: ‘apprenticeship’ (it combine hands-on work with classroom learning to train the apprentice).” (Group 942)

“In this video we can see a not too complex morphology in the word - formation processes. The language use is quite simple.” (Group 946)

“Since the beginning of the interview, we can highlight the semantic field of the school. Lorraine Tonks uses acronyms to describe her Educational Journey, like: C.P.D, SATs, FAMAS. In this type of word-formation, the contraction delivers initial letters that constitute well-formed words about specific professions.” (Group 940)

The discussion on the results of RQ2 helps in understanding RQ3 (“To what extent can the educational use of corpus construction and annotation foster students’ textual awareness”), as most of the students have indirectly commented on textual genres by making reference to specific genres and the competences required to understand them:

“Then, our favourite part occurred: the correction of the auto-generated subtitles provided by YouTube. This step required a good listening ability, also because the speaker was Irish and had a really strong accent that sometimes was difficult to understand.” (Group 922)

“What we are proud of is having learned not to stop only at the external aspects but to grasp the wholeness and depth that often hides and therefore have carried out a complete analysis also from the textual, lexical and morphological point of view (hoping to have succeeded). We are proud and satisfied with our work and also with the choice of video.” (Group 923)

“Furthermore, in our personal opinion, the video was really interesting under many aspects; it was useful not only for the technical purpose of this project, but it was also fascinating because of the topic itself. It was quite recent, and it dealt with the matter that, as students, deeply interested us: that of education.” (Group 921)

They also commented extensively on some specific genre-related components of the oral and written discourse, revealing an appropriate tendency in recognizing the differences in the use of markers of orality:

“We can also notice a large use of “discourse markers”, for example: “at first sight” used to introduce the school, and “so” and “right” used to start a further oral discourse. There are elements of spontaneous conversation in which we can find words that indicate hesitation, for example: “ehm, ahm”, that are words not planned in the discourse but said spontaneously.” (Group 929)

“Each step was carried out in order to know a certain topic and most were aimed at detecting various grammatical structures that differ in the transition from written to spoken English.” (Group 944)

Some comments also disclosed an appreciation of what was being learnt in an intercultural perspective:

“we can see a Somali proverb: “AMA WAA LA MUUQDAA, AMA WAA LA MAQNAADA.” which translated is “Either be visible or be absent. [...] These girls are an inspirational role model for the younger generation.” (Group 923)

“The continuous comparison with the English school has allowed us to detect the differences that make each school system unique and for this cultural expansion we have firmly considered the project very interesting.” (Group 948)

As for RQ4 (“To what extent does the educational use of videos foster multimodal and digital competences”), students related their improvement especially in terms of teamwork:

“It has surely been the first time for everyone of our group being a part of an analysis of this kind. It has been a long-lasting work that took us some time and surely we had our difficulties. For example, we can mention the days we had to complete the Visual Sequencing column for the video because we had to change almost completely what we had written as a result of an unfortunate misunderstanding and that caused us to get stressed out.” (Group 923)

“This project was an excellent initiative because it allowed us to be close even though we were far away because of the pandemic. In a period as hard as the one we are going through, to work in a group made each of us feel less alone and helped us overcome with a little more serenity this period. The live interaction with colleagues that lacks in this period has been substituted by the online group work that has made us understand even more how important it is to work in a group and know how to organize with others. This project also allowed us to expand our knowledge in the field of technology using programs that we did not use often before. The group spirit and the complicity that has been established between us has ensured the realization of this project that is now coming to an end. We are fully satisfied with this work and believe that cooperation between colleagues is essential and we hope that there will be other projects like this in the future.” (Group 928)

“This project gave us the possibility to work with multimodality, and the way in which it perfectly describes and explains the topic through music, sounds, signage and language.” (Group 929)

“We have developed IT technological skills” (Group 923)

“This project has also enabled us to better consider the importance of teamwork. And it is precisely this cooperation that has made our work less burdensome and less complex in its entire course.” (Group 939)

The last comment makes it possible to stress the importance of collaborative work. One of the weaknesses of online teaching is the type of relationship it is creating. Systems such as Zoom, Teams etc. tend to encourage rather old classroom models in which the teacher gives a lecture and the students listen to it. In other words, the classic situation where 21st technology is implementing 19th century educational classroom teaching at any level – whether in the virtual or in presence type – must stimulate dialogue between the students/peers. The most obvious way to do it is via this kind of project work. Video corpora are obvious possibilities for peer interaction in that they allow students to collaborate in the selection, transcription, annotation and critical analysis of a set of videos on an agreed thematic.

Already within this structure, we can see the principle of stage-by-stage project work which is in the current author's opinion, the key point in fulfilling a specific object to the satisfaction of all the students and teachers involved. The previous sections have described the various stages, it is obvious that some of the stages can be omitted and others can be highlighted or introduced. It is also clear that the work carried out by a single group can be redeveloped by other groups of students on different occasions so that the whole procedure relies on the value that can be added to previous stages in a way that millennials can appreciate and understand. The function of the step-by-step model is to suggest to students how to think in analytical terms and how to communicate with each other in terms of the stage achieved. In other words, a description of a result becomes the first step in the communication to each other about the next stage to be achieved. If the individual student is not clear about the framework, s/he cannot communicate to other students. Providing a framework of stages provides a key tool which is missing in most online teaching. The current paper has described certain stages and results have shown 95% of the students have completed all the stages. This high percentage is a sign of the value of group work which combines with the success in achieving the language course objectives.

References

- Ackerley, K. and F. Coccetta (2007). Enriching Language Learning through a Multimedia Corpus. *ReCALL* 19(3): 351–370.
- Baldry, A. and D. Kantz (2022). Corpus-assisted Approaches to Online Multimodal Discourse Analysis of Videos. In *Analyzing Multimodality in Specialized Discourse Settings. Innovative Research Methods and Applications*, ed. by V. Bonsignori, B. Crawford Camiciottoli and D. Filmer, 1–22. Vernon Press: Wilmington.
- Baldry, A. and P. Thibault (2006). *Multimodal Transcription and Text Analysis. A Multimedia Toolkit with Associated On-line Course*. London and Oakville: Equinox.
- Baldry, A. and P. Thibault (2020). A Model for Multimodal Corpus Construction of Video Genres: Phasal Analysis and Resource Integration Using OpenMWS. In *Multiliteracy Advances and Multimodal Challenges in ELT Environments*, ed. by N. Vasta, and A. Baldry, 159–173. Udine: Forum Editrice.
- Baldry, A., Coccetta F., and D. Kantz (2022). What if? Healthcare Simulations, Online Searchable Video Corpora and Formulating Hypotheses. In *Analyzing Health Discourse in Digital Environments*, ed. by A. F. Plastina, 126–146. Newcastle upon Tyne: Cambridge Scholars.
- Baldry, A., D. Kantz, A. Loiacono, I. Marenzi, D. Taibi and Francesca Tursi (2020). "THE MWSWEB PROJECT: Accessing medical discourse in video-hosting websites." *Lingue e Linguaggi*, 40, 433–472.
- Bernardini, S. (2003). Designing a corpus for translation and language teaching: The CEXI experience. *TESOL Quarterly*, 37(3), 528–537.
- Böhme, G. (2012). *Invasive technification. Critical essays in the philosophy of technology*. London: Bloomsbury.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393. doi:10.1111/lang.12224

- Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153–172. doi:10.1080/09588220902778328
- Cambria, M., C. Arizzi, and F. Coccetta (eds.) (2012). *Web Genres and Web Tools*. Como: Ibis.
- De Wilde, V., M. Brysbaert and J. Eyckmans (2019). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition*, 1–15. doi: 10.1017/S1366728918001062.
- Flowerdew, L. (2015). Data-driven learning and language learning theories. In *Multiple affordance of language corpora for data driven learning* ed. by. A. Leńko-Szymańska and A. Boulton, 15–36. Amsterdam: John Benjamins.
- Frankenberg-García, A. (2012). Raising teachers' awareness of corpora. *Language Teaching*, 45(4), 475–489. doi:10.1017/S0261444810000480
- Frankenberg-García, A., L. Flowerdew and G. Aston (eds) (2011). *New Trends in Corpora and Language Learning*. London: Bloomsbury.
- Granger, S. (2015). The contribution of learner corpora to reference and instructional materials design. In *The Cambridge handbook of learner corpus research*, ed. by S. Granger, G. Gilquin and F. Meunier, 486–510. Cambridge: Cambridge University Press.
- Halliday, M.A.K. and R. Hasan (1976). *Cohesion in English*. London: Arnold.
- Jablunkai, R. and E. Csomay (eds.) (2022). *The Routledge Handbook of Corpora in English Language Teaching and Learning*. London: Routledge.
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal*, 4, 1–16.
- Liu, D. and P. Jiang (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, 93(1), 61–78. doi:10.1111/j.1540-4781.2009.00828.x
- Maribel Montero Perez, M. and Michael P.H. Rodgers, (2019). Video and language learning. *The Language Learning Journal*, 47(4), 403–406. doi: 10.1080/09571736.2019.1629099
- McEnery, T. and A. Wilson (1997). Taching and language corpora. *ReCALL*, 9(1), 5–14. doi: 10.1017/S0958344000004572
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Perego, E. and C. Taylor (eds.) (2022). *The Routledge Handbook of Audio Description*. London and New York: Routledge.
- Poole R. (2022). “Corpus can be tricky”: revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 35(7), 1620–1641. doi: 10.1080/09588221.2020.1825095
- Poole, R. (2018). *A guide to using corpora for English language learners*. Edinburgh: Edinburgh University Press.
- Sindoni, M. G., E. Adami, S. Karatza, I. Marenzi, I. Moschini, and S. Petroni (2019). *The Common Framework of Reference for Intercultural Digital Literacies*. <https://www.eumade4ll.eu/wp-content/uploads/2020/02/cfridil-framework-linked-fin1.pdf>
- Taibi, D. (2020). Analysis D: Shaping Digital Identities through Group Study: A Simulated Case Study Using Online Tools and Videos to Explore Animal-human Interactions. In *Multiliteracy Advances and Multimodal Challenges in ELT Environments* ed. by N. Vasta and A. Baldry, 190–195. Udine: Forum.
- Vasta N. (1996). *Università e Lingue alle Elementari. La formazione degli insegnanti elementari in un centro linguistico universitario nel quadro del Progetto DIRELEM-LIJE*. Udine: Campanotto.
- Vasta, N. and A. Baldry (eds) (2020). *Multiliteracy Advances and Multimodal Challenges in ELT Environments*. Udine: Forum.

Authors

Mariavita Cambria, PhD

Associate Professor, University of Messina, Department of Ancient and Modern Civilizations, via Panoramica dello stretto 480, Eden Park 98168 Messina, Italy, e- mail: mariavitacambria@gmail.com, mcambria@unime.it

Izredna profesorica, Univerza v Messini, Oddelek za starodavne in sodobne civilizacije, via Panoramica dello stretto 480, Eden Park 98168 Messina, Italija, e- pošta: mariavitacambria@gmail.com, mcambria@unime.it