

## Using Digital Content to Enhance Education: VLogs as a Resource for MA Theses in Linguistics

Mădălina Chitez<sup>1</sup>

### Abstract

*Much of the research dealing with digital technology in university education has emphasized its impact on students' engagement with academic studies. The use of digital content in applied linguistics also has a high didactic potential, considering that one of the defining features of this field is the connection it creates between linguistic observations, the resulting analysis and the real world. In the case of corpus linguistic approaches, applications are self-directing due to the data-oriented character of the discipline. However, teaching applied corpus linguistics can be a challenge to university teachers, who have to tackle issues of motivation, methodology and data interpretation.*

*Drawing on technology-supported education, the paper argues that students can be successfully motivated to write academic papers in linguistics if they elaborate on topics related to real-life interests specific to their peer group. These topics imply the use of data resources such as vlogs, blogs, forums or video platforms. I will discuss several digital content solutions that a MA thesis can adopt so that students can resonate with their learning and academic tasks. In addition, I will present and debate on the development of the topic for a master's thesis in the area of applied corpus linguistics. Thesis supervision challenges and recommendations will also be integrated into the discussion.*

**Keywords:** *applied linguistics, corpus linguistics, new technologies, digital content, MA thesis, thesis supervision, academic writing, innovative teaching, teaching linguistics, vlogging;*

### 1. Introduction

Digital technology has taken over all areas of modern life, being both imperative and overwhelming. In higher education, this trend is particularly noticeable as numerous teaching methodologies, support instruments and materials encourage the use of new media and web products. This, in exchange, guarantees that the learning content is still appealing to the nowadays student who is used to multiple app use. Studies [1] mention that students evaluate positively digital practices that allow them to "organize and manage the logistics of studies" (e.g. Moodle), to be flexible (e.g. library databases) or to see information in different ways (e.g. videos outside university). It appears that the benefits of using technology in education are much dependent on the purpose of the activity (communication, information, supplementary learning material) and the student target group (age or specialisation). It is generally agreed that digital practices need to be embedded in tested methodologies so that the intended teaching unit should be effective.

In humanities, social sciences and linguistics, HE tutors and teachers are also extremely motivated to make use of the digital world's advantages: availability, variability, accessibility, authenticity and, not least, "coolness". They acknowledge the importance of keeping in line with students' expectations and preferences for activities with a highly interactive and mobile character. Implementing technology in the linguistics course either serves the communicative-informative purposes briefly mentioned before or the digital content is the centre of the study activity. By switching the focus on the content, the technology is not merely an instrument but a research database. For example, if a YouTube video could be an interesting modality to explain the functionality of a device in engineering, it could also become a relevant linguistic sample for the student interested in specialised language.

### 2. Digital content in teaching linguistics

#### 2.1 Context

During most bachelor study programs, the discipline of linguistics is embedded into standard humanities study programs (i.e. modern language, language communication) and it is based on the traditional paradigm of grammar rule theory and practice. At the master's level, in the majority of programs, a shift can be witnessed away from the simplified view of the discipline towards more applicable topics such as specialised language, corpus linguistics, translation or language acquisition

---

<sup>1</sup> West University of Timișoara, Romania



studies. In this context, technology becomes the ideal means of extracting and inventorying the latest language developments which could be of interest to the student.

### 2.1 Internet linguistics: example of a guide

When Internet started to be a popular reality, established linguists recognised its potential for teaching and research. David Crystal's book *Internet Linguistics: A Student Guide* [2] is a relevant study on Internet materials can be useful for the field of linguistics. Even if the author rules out the non-textual outputs of the Internet (videos, images) by primarily focusing on the written language on the Internet, the chapters provide several interesting language-related case studies: Twitter, language change, multilingualism and applied Internet linguistics.

### 2.2 Applications in corpus linguistics

Corpus linguistics is a special branch of linguistics which has gained visibility and showed applicability as a package together with computational linguistics. Since corpora are computer-supported collections of texts or transcriptions of spoken material, linguistic studies that use corpora as data pool are technology-related. There are numerous studies which describe the usefulness of web products for corpus linguistics. By including an array of case studies, Hundt et al. [3] argue that "the web can be used as a corpus itself" or it can serve as a source for large corpora (p. 2), thus clarifying the concepts of *Web as Corpus* (WaC) versus *Web for Corpus* (WfC).

Another area where corpus linguistics and technology join forces for the benefit of educational practices is Computer-Assisted Language Learning (CALL). The study of Chitez et al. [4], among others, elaborates on the solution of developing and implementing an academic writing support platform, *Thesis Writer*, which includes an interactive corpus-based linguistic tool. The tool can be used by students in real time while they are writing their bachelor's or master's thesis.

## 3. Master's thesis topic in linguistics based on digital content

There is abundant digital content available online which can be used by students for their master's theses. Good web-supported thesis topics can be identified by answering questions related to the students' observations: *Which are the prominent linguistic features of online newspaper articles? Are there any typical translation mistakes in movie subtitles? Which are the favourite discussion topics on a particular YouTube channel? Is there a typical blog text? How can vloggers improve their visibility by using better language?*

### 3.1 A case study

During her final year of study, a master's student enrolled in the study program British and American Studies (code: AD) decided to use the corpus linguistics methodology in her thesis. AD was also very determined to answer one research question she was particularly interested in: *What are the linguistic characteristics of the speech used by the best lifestyle vloggers?* The intention was to evaluate if there is any correlation between typical linguistic features and their success with the online audience.

### 3.2 Why vlogs?

For young people, vlogs are the ultimate source of entertainment. In their free time, students watch dozens of videos created by popular vloggers. A vlog is a hybrid digital product resulting from the transformation of written blogs into video materials. Vlogs are also defined as "video collections that serve both as an audiovisual life documentary, and as a vehicle for communication and interaction on the Internet" [5].

### 3.3 Data analysis stages

The preliminary step in corpus-based studies is the creation of the digital database. The following procedures were performed which helped the student organise and prioritize research tasks:

- a. *Data representativeness*. A discussion was held between the student and the thesis supervisor regarding the types and selection of vlogs to be included in the case-study corpus. The topic of the vlogs had to be consistent (i.e. lifestyle / makeup). Since online popularity was a central aspect of the research question triggering the analysis, selection criteria had to be identified: number of views and followers and/or beauty vlogger ranking lists. Similarly, a solution was reached by



making the corpus gender-balanced, comparable in point of English variety represented (British English versus American English) and time-balanced (not older than 5 years).

<i>The US corpus:</i>		
Carli Bybel	<a href="https://www.youtube.com/watch?v=lruXBzjMzsM">https://www.youtube.com/watch?v=lruXBzjMzsM</a>	7.58 min
Jefree Star	<a href="https://www.youtube.com/watch?v=FB2UE_SWIXY">https://www.youtube.com/watch?v=FB2UE_SWIXY</a>	19.41 min
Jacklyn Hill	<a href="https://www.youtube.com/watch?v=Y4fur_12EEc&amp;t=383s">https://www.youtube.com/watch?v=Y4fur_12EEc&amp;t=383s</a>	11.05 min
Kathleen Lights	<a href="https://www.youtube.com/watch?v=ePmnAZUQIkM">https://www.youtube.com/watch?v=ePmnAZUQIkM</a>	8.31 min
Manny MUA	<a href="https://www.youtube.com/watch?v=cHMb0qH-Xbw">https://www.youtube.com/watch?v=cHMb0qH-Xbw</a>	10.1 min
James Charles	<a href="https://www.youtube.com/watch?v=9u1kqc0dT4M">https://www.youtube.com/watch?v=9u1kqc0dT4M</a>	15 min
<i>The UK corpus:</i>		
Wayne Goss	<a href="https://www.youtube.com/watch?v=bBLHE0ME-GA&amp;t=79s">https://www.youtube.com/watch?v=bBLHE0ME-GA&amp;t=79s</a>	6.05 min
Lisa Eldridge	<a href="https://www.youtube.com/watch?v=f-SZA1tkViU&amp;t=460s">https://www.youtube.com/watch?v=f-SZA1tkViU&amp;t=460s</a>	13.50 min
Looking for lewys	<a href="https://www.youtube.com/watch?v=Za1mXA9WcC0&amp;t=419s">https://www.youtube.com/watch?v=Za1mXA9WcC0&amp;t=419s</a>	9.14 min
Zoella	<a href="https://www.youtube.com/watch?v=CrOAVLICvhw&amp;t=110s">https://www.youtube.com/watch?v=CrOAVLICvhw&amp;t=110s</a>	10.30 min
Kaushal Beauty	<a href="https://www.youtube.com/watch?v=oKQbHNR6YtQ&amp;t=1s">https://www.youtube.com/watch?v=oKQbHNR6YtQ&amp;t=1s</a>	8.49 min
John Maclean	<a href="https://www.youtube.com/watch?v=MyEqfBkzESU&amp;t=806s">https://www.youtube.com/watch?v=MyEqfBkzESU&amp;t=806s</a>	17.39 min

Table 1: Vlog corpus (AD selection)

- b. *Data processing.* The selected vlog episodes are in YouTube format. The next challenge was to transcribe the monologues in the videos. AD proposed the use of the YouTube-integrated automatic speech recognition technology. She noticed that this technology is not 100% accurate due to the different pronunciations, accents, speech pace and tonality of the vloggers. Another inaccuracy issue concerned punctuation marks, which were omitted during automatic transcriptions. AD had to refine the transcriptions manually, while listening to the videos, by correcting false recognitions and punctuation marks. The time invested in this procedure varied: the manual correction of transcriptions, for each video, took between 3 hours and 4 hours depending on the duration of the video, the variety of English used and the speech typology. The whole corpus (12 videos) required around 50 processing hours.

<Vlog-001.1>

Hey, guys! welcome back to my channel. <chuckle> Today's video is going to be a testing of a new foundation that I saw online, as a couple people were tweeting me about this foundation saying Hey man, you should try this out, it looks a little funky it's a little weird and I'm like, okay I'm totally down to try out and test out some products for you guys, as long as you guys are requesting it. And I did see a BUNCH of you guys requested me to review this foundation and so I bought a couple shades, this is the foundation right here.

(source: Manny MUA / <https://www.youtube.com/watch?v=cHMb0qH-Xbw>)

Table 2: Excerpt from vlog transcription

- c. *Data-driven analysis.* Since the topic is highly exploratory, the student had to go over transcriptions in a “detective-like” manner. This implied that she had to use her observations on the typical language usage patterns in the vlogging sessions and determine whether these patterns appear in correlation with other linguistic features or whether they display noteworthy frequencies. AD replicated some of Riboni’s analyses [6] by focusing on rhetoric move linguistic markers (e.g. *Hey, guys!*), formulaic expressions (e.g. *I did see a bunch of you guys*) and prominent linguistic patterns.

## 4. Discussion

### 4.1 Supervisor’s observations

It is possible to involve students in technology-related topics that are connected with their life outside university either by allowing them to select a theme that reflects their observations or by offering them a list of topics which could be compatible with their own interests or hobbies.

As regards the methodology of corpus linguistics, it is advisable that the student has had previous minimum exposure to it (e.g. an introductory course, training sessions). Otherwise the supervisor will



have to address too many challenges at once and the students might be discouraged in the process. If the student knows how corpora function, then the whole analysis-stage guidance is stimulating.

#### 4.2 Conclusions

Working with students on topics that are particularly interesting to them is exciting and pedagogically rewarding. By securing motivation during the master's thesis writing, the rest of the supervising stages seem less challenging. Technology-related topics or digital content research in linguistics is certainly a path that university teachers could envision in order to be able to demonstrate to their students that study subjects have real-life applicability.

#### References

- [1] Henderson, M., Selwyn, N., and Aston, R. "What works and why? Student perceptions of 'useful' digital technology in university teaching and learning", *Studies in Higher Education*, 2017, 42: 8, 1567-1579, DOI: 10.1080/03075079.2015.1007946.
- [2] Crystal, D. "Internet linguistics: a student guide". London, Routledge, 2011.
- [3] Hundt, M., Nesselhauf, N., and Biewer, C.(eds). "Corpus linguistics and the web", Amsterdam/New York, Rodopi, 2007.
- [4] Chitez. M. "Corpus-based phraseology use within an academic writing platform: a plagiarism check", *Conference Proceedings. ICT for Language Learning. 9th edition, Florence, Italy, 17 -18 November 2016*, Pixel, Libreriauniversitaria.it, 2016.
- [5] Biel, J.I. and Gatica-Perez, D. "Voices of Vlogging", *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, D.C., May 23-26, 2010*, The AAAI Press, Menlo Park, California, 2010, 211-214.
- [6] Riboni, G. "THE YOUTUBE MAKEUP TUTORIAL VIDEO: A preliminary linguistic analysis of the language of "makeup gurus"", *Lingue e Linguaggi*, 2017, Vol. 21, 189-205.  
Internet linguistics: A student guide (New York, Routledge. 2011)