Didactic Modeling of the Videotutorial of Language Teaching Multimodal Text Genre: a Preliminary Study

Joaquim **DOLZ-MESTRE***
Aina **MONFERRER-PALMER****

* Doctoral Degree in Language Teaching from the Université de Genève - UNIGE (1991). Honorary Professor at the Department of Psychology and Educational Sciences at the Université de Genève - UNIGE. Contact: Joaquim.Dolz-Mestre@unige.ch

** Doctoral Degree in Philology and European Cultures from Universitat Jaume I - UJI (2015). Assistant Professor at the Department of Didactics of Language and Literature at the Universitat de València - UV. Contact: aina.monferrer@uv.es

Summary:

This study analyzes the use of Videotutorials for Language Teaching (VLTs) by language learners through multimodal analysis of a corpus and through surveys of VLT users. Six types of VLTs have been identified and evaluated which ones are most valued by language learners. The results show that VLTs of specific tips for solving errors and learning obstacles are the most highly rated. In addition, VLTs showing the resolution of official tests and the video correction of written expressions are also highly valued by their users. Learners use VLTs as a model for understanding and practicing language structures. Observed practice has been shown to aid in repetition training and progressive self-correction. Most respondents prioritize the usefulness of the content over the technical quality of the VLTs. In summary, the results suggest that VLTs are a complementary tool valued by language learners that allows them to improve their autonomous learning and progress in their language skills.

Keywords:

Videotutorial for Language teaching. Hypergenre. Didactic sequence.

Didactic Modeling of the Videotutorial of Language Teaching Multimodal Text Genre: a Preliminary Study

Joaquim Dolz-Mestre; Aina Monferrer-Palmer

1. Introduction

This study, situated in the line of sociodiscursive interactionism (ISD), considers the potential of the videotutorial in language teaching. At this point, from the discourse theories, there is a lack of characterization of the videotutorial in language teaching as a text genre. This, added to the complexity of being a multimodal text, has caused that the approach to the videotutorial for language teaching (hereinafter, VLT) as a didactic tool in digital environments is still shallow.

The videotutorial is a genre of text that evolved from the instruction manual, but also from tutorials in technical learning, whose generalization on the web is still recent. We ask ourselves the following questions: What are the specificities of videotutorials for teaching, how are they used by teachers, and what particular uses do language learners make of them? Regarding the use of videotutorials among language learners, their function, values and preferences are unknown beyond a series of ideas with little empirical foundation in relation to the duration and technical perfection of the product. A study on the use of VLTs such as this one will allow us to begin to analyze their didactic potential as is already being done with similar genres such as podcasting (OLIVEIRA; DIAS, 2021).

Once a bibliographical review of the videotutorial has been carried out, starting from studies that reflect on its birth and evolution up to specific studies on the use of videotutorials in language teaching, the research has three objectives:

1. To start from the characterizations of the videotutorial genre of language teaching in order to propose a consistent didactic model that presents the most important dimensions for language learning.
2. To analyze the preferences of users of videotutorials for language learning, both in the choice of platform and in the use according to needs.
3. Based on the data obtained, reflect on the possibilities of using videotutorials in language teaching.

2. State of the Art: studies on the videotutorial genre

First, we cite Mogos and Trofin (2015: 38), who start from a categorization of Charaudeau (1997) to distinguish two levels of analysis of videotutorials: *a) the pragmatics of discourse* (discursive modes, macrostructural organization), and *b) the textual and enunciative configuration* (structure and dimensions of the text, role of enunciation in the text, microstructure, syntax and stylistics). These authors emphasize the subversion of the figure of the expert, since, in the world of videotutorials, *prosumers* are both consumers and producers:

> Traditional media still relies on standard media genres, but the web creates an environment where message producers are not only the professionals invested with authority, who are aware of rules and norms that are governing the media communication process, but the members of the audience, too (MOGOS; TROFIN, 2015, p. 38).

Genres are social constructions that streamline intercomprehension. Heyd (2009, p. 241 apud MOGOS; TROFIN, 2015) defines the videotutorial as a multimodal text *supergenre*; he considers this *supergenre* as a moderately stable fact that has its antecedents in pre-digital discourse. The difficulties of discursive classification in the web and social network context are defined by two phenomena: genre hybridization and individualization. *Supergenre* subcategories are more changeable and adaptable in the digital environment because:

"The Web is fluid, unstable and fast-paced; on the other hand, genres on the Web are instantiated in Web pages, which are a complex type of document, more composite and unpredictable than paper documents" (MOGOS; TROFIN, 2015, p. 39).

A sign of this interesting complexity for socioconstructivism is a feature of videotutorials that traditional tutorials do not have: the ability to interact through comments left by users, star ratings or even communication via chat if the premiere is synchronous, as in YouTube and Twitch live streams.

There are multiple attempts to classify types of videotutorials. For example, Brezeale and Cook (2008 apud MOGOS; TROFIN, 2015, p. 40) identify four groups of methods to classify videos: *text-based approaches*, *audio-based approaches*, *visual-based approaches* and possible combinations of these three variables. From a pragmatic point of view, videotutorials, with their strong instructional intentionality, are also known as *how-to-do*. According to these authors, Mogos and Trofin (2015, p. 42), searching YouTube for "how to" produces 155 million results; and searching for *tutorial* produces 83 million. The most viewed YouTube tutorials are those on *beauty tricks*. In contrast, the most viewed YouTube tutorial in 2015 was "10 *how to* pranks for home and office". However, the most viewed videos on YouTube according to global data are pop and children's videoclips, a fact that also tells us the worrying number of hours that young children are plugged into this type of screen-mediated entertainment. Language learning videotutorials are in the tenth position of the most viewed types of tutorials on YouTube.

As we pointed out in *Docentes hiperdiscursivos* (MONFERRER-PALMER, 2022), the *youtubization* of television formats has accelerated since the confinement of 2020. The general viewer, and especially the young, has become accustomed to consuming amateur content, with clearly homemade video and audio editing, produced by what is known as proactive consumer or *prosumer*. In this context, teachers who are experts in their subject, but not necessarily in audiovisual editing, have launched into the publication of videotutorials that, although they look like home editing, offer consistent knowledge typical of a teacher who is an expert in a subject and with remarkable informative and didactic skills. Mogos and Trofin (2015, p. 42) state that the number of views of a video on the Internet is more related to the content of what is explained than to its aesthetic perfection. Furthermore, they state that most users do not watch the entire clip but only until they find an answer to their question.

Internet rules do not apply to traditional television and vice versa. According to the same authors, Mogos and Trofin (2015, p. 43), the most important thing about a videotutorial on the Internet is its usefulness. Despite the fact that, as mentioned, the sites where most videotutorials are located today are private multinationals, the idea of sharing videos altruistically to help unknown people means that the spirit of *prosumers* does not cease to be that of sharing and helping in a context of freedom of knowledge transmission. Tarquini and McDorman (2019, p. 148) define videotutorial as "asynchronous instructional videos (often screencasts, but also animation or live-action footage) providing step-by-step guidance for specialized activities". In "Content assimilation and learning through the use of videotutorial", Tapia, Delgado and Garcia (2020, p. 2) state that:

> Videotutorials are multimedia elements that allow access to the informative part of the teaching process in a dynamic way that attracts the learner to follow them as a guide in some process. The videotutorials are very well seen since they allow to review the content as many times as necessary until the student achieves the desired knowledge or the development of some proposed skill.

Tarquini and McDorman (2019, p. 146), performing a formal analysis of the phenomenon of videotutorials, highlight that these have been exploited by technology companies as a complement to technical

documentation. In other words, they have become an audiovisual version of instruction manuals. Similarly, videotutorials for language learning are an audiovisual version of the written explanations of regulations in textbooks.

These authors Tarquini and McDorman (2019, p. 147) distinguish three types of videotutorials: academic, corporate and user-made, and take as their corpus of analysis a collection of videotutorials on the use of PowerPoint posted on YouTube. They will analyze videotutorials as a semiotic artifact conceived as an emerging textual practice in the current landscape of the learning society. Videotutorials are a unique multimodal form of instructional text. Some authors consider the *screencast* as a partial synonym of the videotutorial for being the prototypical form of the videotutorial with *software* explanations. Terms associated with the videotutorial would be *online tutorials*, *electronic tutorials*, *interactive tutorials* or *cumputer-based tutorials*, among others.

Tarquini and McDorman (2019, p. 147) state that, although it originally referred to the accompaniment of a learner by a tutor, the focus of the *tutorial* concept has expanded to include a variety of instructional sources and tools for individual tutoring of learners, to the point of being able to replace human tutoring. Although it is not being claimed that this is the ideal way to learn, many people try to learn any subject (including languages) with videotutorials and some of them do it successfully.

Tarquini and McDorman (2019, p. 148) have verified the existence of an extensive literature regarding videotutorials in teaching, whose effectiveness has been proven in university subjects such as physics or engineering. Some of the benefits that stand out are the adaptation to different learning styles, total time flexibility, complete availability and free of charge. On the other hand, these authors report disadvantages in the use of videotutorials for learning, such as the fact that they are not particularly useful if they are reproduced during classroom time.

These authors, Tarquini and McDorman (2019, p. 149) go back to the Second World War as the first moment in which the videotutorial was used. In addition, they relate it to MOOCs and *flipped classrooms*. They consider it an undeniable fact that, at present, videotutorials are having a profound impact on teaching and learning methodologies, an aspect that this article considers in the specific framework of language learning.

It has been demonstrated that the combination of oral and visual materials reinforces the stimulation of memory and cognition thanks to the activation of two separate but interconnected neurological systems (verbal and non-verbal in images). This is known as *Dual Coding Theory* (PAIVIO, 1971).

Tarquini and McDorman (2019, p. 151) pick up the study conducted by Gromik (2007) "Video tutorials. Camtasia in the EFL classroom" as the first attempt to evaluate the impact of VLTs on foreign language learners. In this first research with Japanese learners of English who were engineering students, shortcomings were detected in the EFL method based on videotutorials in terms of learning oral skills and vocabulary acquisition.

Mogos and Trofin (2015, p. 44) distinguish two ways of approaching the videotutorial: the professional and the homemade, and compare them. The main differences they find are *a)* that the professional approach involves all the rules and conventions of video editing and television production (HD, adequate lighting for each shot, many details and well cared for, action carried out by actors and also a good narrative established by a professional script), while *b)* homemade videotutorials do not follow these conventions. On the other hand, the main similarities between professional and amateur videotutorials are *a)* the instructional intentionality, and *b)* the free accessibility through massive internet platforms.

In this article, videotutorial values such as editing quality, the professionalism or experience of the speaker and the author (who are often the same person in non-professional videotutorials), the duration and the platform on which they can be found have been taken into account. These items have been asked about both in the VLT user surveys and when proposing a didactic modeling. In order to carry out an effective analysis and categorization of language videotutorials, it is important to consider that different issues are intertwined in them, such as experience in the teaching subject, communicative skills and audiovisual editing skills, with the extra complexity demanded by metaverbal reflection.

In amateur videotutorials, Tarquini and McDorman (2019, p. 151) observe a user addressing other users in an informal learning context, with videos that are not usually scripted or carefully post-produced and in which the verbal texture usually contains lack of cohesion and coherence, some performative instability, spontaneous speech and hesitation. The tripartite classification between *academic*, *corporate* and *user-generated* videotutorials offered by these authors is only indicative and does not avoid overlaps between categories.

According to these authors, Tarquini and McDorman (2019, p. 152), academic videotutorials must be designed as part of a broader curriculum and are the subject of activities, assessment and interaction with students on a face-to-face, hybrid or fully distance teaching campus. They are intended for a more or less homogeneous audience with common objectives and competencies. Moreover, they claim that corporate videotutorials tend to have a higher editing quality (TARQUINI; MCDORMAN, 2019, p. 154) than user-generated videotutorials (made from *do-it-yourself*). According to Omizo (2012), Internet videotutorials are made in the context of an altruistic knowledge-sharing culture on the Internet. Tarquini and McDorman have also observed that user-generated videotutorials in the altruistic landscape of the learning society are usually made in languages with fewer users, while corporate ones are usually in English and subtitled due to their commercial profitability purpose.

Regarding the object of study of this research, in VLTs it is difficult to distinguish between *academic tutorials* and *user-generated tutorials*. In the case of VLTs in profiles such as *Learn* [English] *with* [Marisa], language teaching expertise is confused with that of a native speaker or advanced self-learner of an L2 or L3. In reality, being a speaker does not imply being an expert in language teaching, not even in the norms of the standard model, just as having a body does not imply being a doctor. Therefore, it seems logical to think that it is more reliable to be able to identify the person behind the VLT channel in order to verify his or her training and experience than to blindly rely on channels with ambiguous and generic names (such as *Learn* [language], *Easy* [language] or similar) if, in addition, they are not endorsed by any prestigious academic or normative institution.

In *Docentes hiperdiscursivos* (MONFERRER-PALMER, 2022), it is stated that the language teacher who creates videotutorials in his spare time and publishes them in the open for the love of art (while helping his official students), is required to have experience in his job of making the interstices of the language understood, but he cannot be required to have great knowledge of audiovisual editing. As in corporate videotutorials, in academic or user-generated VLTs it is common to find self-promotion if the author has achieved a minimum of success and is creating paid materials or courses in parallel to what he offers for free.

In addition, it is inevitable that the platform conditions the formal characteristics of the VLT. There is a contradiction that plays against the complexity of the communicative and linguistic phenomenon: there are platforms that increasingly demand shorter videos and this leads to the fact that information must be subdivided into micropills. This makes it difficult to explain complex and nuanced concepts such as the genuineness of reference mechanisms, lexical richness, the use of punctuation marks or semantic evolution in a text.

So if about 100 videos lasting 15-20 minutes would be needed to explain the essential rules of a language such as Catalan at level C1 (as has been done in the VLT project De mica en mica), based on 1 or 2 minute videos, a thousand would be needed, and the common thread would probably be lost when viewed from a global planning of language courses according to the levels established by the *Common European Framework of Reference* (COUNCIL OF EUROPE, 2002). Tarquini and McDorman (2019, p. 166) underline the relationship of the semiotic practices involved in videotutorials with the *Companion to the Common European Framework of Reference* (COUNCIL OF EUROPE, 2018), which establishes audiovisual skills assessment descriptors (*AV watch*, *AV listen*, *AV read*, *AV speak*, *AV write*, *AV produce*).

From language teaching, Román (2018, p. 193) analyzes the videotutorial as a new discursive genre emerged in the digital society of learning and perceives it as a type of text that is situated in what Mostacero (2004) calls *scripturality* (*escrituralidad*). That is, a text that is neither purely written nor oral, but predominantly oral. Román works on the elaboration of videotutorials on spelling rules, which 3rd ESO students in València elaborate and then share with their 1st ESO classmates.

Román (2018, p. 206) states that, at the end of his didactic sequence for working with the videotutorial text genre, half of the students managed to identify instructional textual sequences and defends the didactic usefulness of the VLT as the final product of a didactic sequence in ESO students (ROMÁN, 2018, p. 206). It can be added that the elaboration of videotutorials with students can serve as an interesting tool to work on metalinguistic reflection in the classroom. In this sense, we highlight the genealogical connection of the VLT in the classroom with the *secuencia didáctica para aprender gramática* (SDG) by Zayas and Camps (2006).

Similar to the initial research on the videotutorial as a didactic tool from the ISD, in the context of Portugal and Brazil, work is being done on the didactic modeling of another genre, or hypergenre (MIRANDA, 2010), of emerging multimodal text with obvious parallels to the VLT and didactic potential for language teaching, such as the podcast (OLIVEIRA; DIAS, 2021).

3. RESEARCH METHODOLOGY

Situated in the theoretical framework of ISD, here the text genre is analyzed in terms of the regularities observed in current practices. In the previous section, the emergence of the VLT was placed in a historical perspective and in relation to other discursive practices. The VLT is intended for a particular sphere of activity: language teaching and learning, and in an environment that, since the pandemic, has become consolidated and widespread (MONFERRER-PALMER, 2022). In this sense, two hypotheses are at the basis of our work:

1. For pragmatic reasons linked to the sphere of the proposed activity, the VLT is a text genre with formal particularities that distinguish it from other types of videotutorials: cooking, software, make-up, sports.
2. On the part of teachers and learners, the use of VLT is chosen as a complement to other teaching modalities, to reinforce autonomous self-learning. Our hypothesis is that language learners who regularly attend courses also use the VLT to extend or reinforce knowledge autonomously.

The methodology contemplates three work stages:

The first stage of the research consisted in analyzing the most used VLT by the new generations. For this purpose, a corpus of 20 VLT (in Catalan, French, English and Portuguese) from the most used social networks for VLT (YouTube, Instagram and TikTok) has been constituted. Each of the videotutorials has been analyzed according to the following criteria: language, number of followers, platform, specific topics within the language teaching channel, visual composition (framing, arrangement of elements, appearance of person or persons, didactic accessories, average time of the videos, whether the person or institution responsible for the channel is identifiable and what professional prestige they have and what type of interaction is proposed).

In the second stage of research, a study is conducted on the practices declared by VLT users through a questionnaire previously elaborated by the researchers. The questionnaire is organized in four axes: *a)* demographic profile, *b)* multilingual profile and learning style, *c)* preference in the consumption of videotutorials, and *d)* preference in the consumption of VLT, and it has been disseminated through two different channels: Instagram stories and Google Forms (the latter, through Twitter, WhatsApp and virtual classroom of students of Magisteri at the Universitat de València). A total of 374 subjects have responded to the survey, of which 200 have done so through stories on the Instagram account of the VLT project De mica en mica (*@ainamonferrer*). In addition, 174 surveys have been collected from students of Magisteri at the Universitat de València, as well as through Twitter and WhatsApp and Telegram groups. The method of processing the data from this survey specifies the frequency of the various choices in a way that allows us to sketch a portrait of the current trends in the use of VLT on the network.

In the third stage, based on the analyses of the previous stages, a didactic modeling of the genre is established and what we consider to be the greatest potential for language learning offered by the VLT is formulated.

4. RESULTS

4.1. Analysis of 20 channels of videotutorials on language teaching (VLT)

The first conclusion from the qualitative analysis of the accounts is that prototypical VLTs are characterized by the presence on screen of a person in close-up. In a dance or fitness tutorial, bodies often appear in close-up so that the viewer can imitate their movements; in makeup videos, the face appears in close-up and the color and lighting nuance is important; in videos of video game players, a close-up of the game screen plus a frame of the player's face in one corner-usually the upper right corner-constitutes a close-up of the game screen[1].

On the other hand, in a VLT it is necessary to observe the mouth, but also the expressiveness of the rest of the face and the gestures of the arms. A common mistake in amateur videotutorials is to add non-diegetic sound in the background to appear more professional, although, according to users' opinions, they often achieve the opposite because the volume of the music below the explanation hinders comprehension.

It has also been observed that, in the Portuguese videotutorial accounts, there is an eagerness to specify those that teach Portuguese from Portugal (possibly a more prestigious variant, as occurs with peninsular Spanish for reasons of historical geopolitics), which we do not find so much in the VLTs elaborated by Brazilians. In the Catalan accounts, there is a clear difference between the VLTs for Catalan as a foreign language (CLE), which are usually accounts that teach standard Eastern Catalan for basic levels, and the accounts that teach Catalan C1-C2 for people who are preparing for the official tests for work purposes, which are usually in standard Western Catalan Valencian due to administrative differences between the Catalan-speaking territories in question (Catalonia and the Valencian Country).

Among the first type, we find the TikTok and Instagram account for learning Catalan with the most followers: Daily Catalan, which teaches CLE with English as the starting language, and in the second, De mica en mica, among others. In the qualitative analysis of the language learning channels (available at this link: *shorturl.at/rsGL2*), a predominance of the Caucasian phenotype and young people of heteronormative appearance as experts was observed, with few exceptions such as the case of Antonio Parlati's TikTok account for learning English, of non-binary appearance.

From the analysis of language learning channels, it has been found that, on YouTube, VLTs tend to be longer. On Instagram, there is a peculiar format which is that of *posts* or entries with several slides and simple graphics in which lists with orientations are presented, vocabulary, phraseology or common mistakes are worked on. In TikTok, VLTs are usually shorter than in YouTube, with homemade graphics and a fast pace, explanations in micro-pill format and with a multiplatform reuse.

The fact is that many creators reuse vertical videos in different social networks. We could include this reuse in what Tarquini and McDorman (2019, p. 165), referring to videotutorials, call *transcreation*. TikTok's vertical and very brief language is penetrating YouTube with YouTube Shorts, and Instagram (with *reels*, promoted in *stories*). Unlike what happens in other types of videotutorials, the vertical format favors the prototypical language videotutorial, in which a person appears, usually in American shot so that the upper limbs and also the face can be seen, essential elements in conventional interpersonal communication.

In TikTok, the most successful VLTs are of the type "*Learn* [language] *with* [own name]". Currently, there is a TikTok account used by many young Spaniards between the ages of 16 and 21 that is *@tuprofesoradelengua*. This is a secondary school teacher of Spanish who is dedicated to correcting errors in the comments written to the same TikTok videos that she posts. This unconventional TikToker, a Valencian teacher in her sixties with more than a million followers and a masterful interaction management, still uses a prototypical element of the classic language videotutorial: the whiteboard.

---

[1] This last provision coincides with that of Twitch *streams*, which was born precisely as a *gaming* platform.

Marked cultural differences have also been found in the videotutorials depending on the source language. For example, those for learning English with Sanskrit or Arabic as the source language often have a male protagonist or no people in the image. Arabic is often taught with videos linked to the Koran. On the other hand, we rarely find non-Caucasian or non-heteronormative people in the most popular Russian, French, Spanish or British English videotutorial channels. After the formal analysis of the corpus of videotutorials, we proceed to outline a proposal for didactic modeling of the VLT:

- Definition: VLT is a multimodal instructional genre intended for self-learning or to reinforce the regulation of language learning in the classroom.
- Communication situation: the VLT is a teaching mediation tool that is proposed in networks in order to facilitate linguistic learning. The sphere of the activity is its learning, which can be school or out-of-school. It contains didactic dimensions, although they are not always formal.
- Thematic content: the object of the VLT is a verbal practice (aimed at comprehension, production or facilitating interactions) or knowledge about languages (rules of use, grammatical rules, literary or sociolinguistic knowledge, among others).
- Structural organization: the stages of carrying out a linguistic procedure predominate, such as *step-by-step* (linguistic competences), although some present or explain linguistic or literary knowledge (source of normative consultation, definitions, categorizations, conceptualizations, among others).
- Textualization of the VLT: although the fine analysis of the linguistic characteristics is still pending, it can be affirmed that in the VLT the marks of enunciative responsibility and all the linguistic elements typical of instructive textual sequences (with exemplification of a practice) predominate with the particularity of an unfolding commenting on the specificities and how to overcome the main obstacles of language learners.
- Multimodal characteristics: the linguistic practice taught is visualized with different oral and written possibilities (different superimposed semiotic systems appear where graphics, images and videos have a fundamental importance as support for the text) reinforcing the features that are the object of teaching and interactive tasks are proposed with the users that involve orality and writing, within the limits of interaction that the material characteristics of the VLT allow.

4.2. Analysis of videotutorial consumption surveys

Some 60% of the 374 respondents are in teaching or are studying to become teachers. Contrary to the claims of some authors such as Mogos and Trofin (2019, p. 42), 72% of the respondents state that they watch the entire videos and not only up to where the specific question they had is solved. Further research is needed on this point to know whether this is a generationally biased or VLT-specific behavioral difference, or whether it is a generalized habit in the consumption of videotutorials.

A majority of people gave more importance to sound than image for VLT consumption. In fact, one of the open answers in *stories* asked "to be able to listen to the video while doing other things", which gives us a clue to a usual method of videotutorial consumption: podcast-type consumption, which also contradicts the widespread prejudice that the shorter a VLT is, the more it is liked, since videotutorials that are consumed as podcasts should be long, minimum about 15 minutes, so that the listener could do another activity at the same time without having to stop and click. Apart from this, respondents attach particular importance to the author of the video being a recognized expert in the subject and little importance to the quality of the image.

Most of the followers on Instagram of the project De mica en mica (*@ainamonferrer*) from which the survey has been disseminated are preparing the C2 of Catalan, and some the C1. In addition, only a minority are simultaneously preparing other language tests (2 French, 2 Portuguese, 1 Danish and 5 English). Only one of the responses is from an Instagram user who is not studying to pass any language level. This denotes the predominance of extrinsic motivation in language learning (especially Catalan in the Valencian Country)

leading to obtaining official qualifications. The vast majority of respondents on Instagram state that they have attended some language course in the last ten years, although most currently state that they are not attending any. Therefore, they are generally users who are preparing for the C1 or C2 Valencian tests mostly on their own (self-study), but with recent experience in learning through formal courses.

As the *Common European Framework of Reference* itself recommends (and for which the European Language Portfolio was created), it is advisable to make language learners reflect on their multilingual learning process as a holistic and *lifelong learning* process (COUNCIL OF EUROPE, 2002, p. 99-120). It is in this context that the questions in this survey would be justified. The main resource for language learning for 49% of the respondents is videotutorials, while for only 24% it is face-to-face classes. Although there is no doubt that a face-to-face course is better than learning through videos (with predictably worse feedback than in a regular, tutored class), we should not underestimate the resource needs of an important sector of the population that either prepares for official language certification tests on their own, or complements their preparation by consulting videotutorials on the Internet.

On the other hand, we must be aware that, to a certain extent, consistent VLT projects on social media channels contribute to the homogenization of the preparation of official certification tests and serve as a reference for teachers when preparing their classes, who in turn will act as VLT prescribers for their students[2]. The issue of videotutorials should be taken seriously if we take into account that data show that more people learn languages through videotutorials than by attending face-to-face classes.

Finally, the results of the survey show that the vast majority of respondents consume tutorials mainly from YouTube, followed by far by the consumption of tutorials on Instagram, and even more distantly, from TikTok. However, if we take into account the survey on the Google form, which was answered mainly by second year primary school students at the University of Valencia, the generation gap becomes evident, since, among the youngest respondents, TikTok is more used for VLT consumption than Instagram (1% more), although the consumption of videotutorials on YouTube is still preeminent in all generations of VLT users.

5. Conclusions

In this last section, based on the modeling of the genre presented above and the description of the main results on uses, the objectives and hypotheses will be reviewed. Is the videotutorial genre being fully exploited? We have observed that, of the six types of language videotutorials that can be categorized according to the analysis of the corpus of 20 VLT channels (see the first subsection of the Results), the one most highly rated by respondents is that of tips specifically targeting errors and learning obstacles. This suggests that language learners who consume them have an eminently extrinsic motivation in learning centered on various personal goals (obtaining a degree, understanding a grammatical problem, organizing a specific text genre, specific phonological training, enriching a lexical field of specialty, or others).

The annotated resolution of official tests and the video-correction of written expressions of real students was the second best rated type of VLT. This can be interpreted to mean that autonomous language learners value mimesis as a useful learning mechanism. It seems that watching a teacher solving oral and written comprehensions, expressions and linguistic mediations, while explaining the steps to follow in the process of successfully solving the task (*step-by-step*), activates the mirror function in the learners and helps them to digest the process of effectively solving the different parts of the tests. The same applies to the annotated video-correction of written or oral expressions of other learners.

We emphasize that learners primarily use VLTs as a model for understanding or practicing novelties of use, as a reflection of observed interactions. The observed practice serves as a model, often developed step by step, and the comments accompanying the practice help a conscious control of the behavior that

---

[2] There is evidence that the videotutorials of De mica en mica even serve as a guide for teachers of preparation courses for the C levels of Valencian for adults.

regulates the verbal activity (VYGOTSKI, 2019). In addition, this self-reflective process allows for repetition training in which the learner can progress by self-correcting step by step. It is an interactive practice that can be done accompanied or alone, despite the fact that doing it alone has its advantages because of the intimacy that helps to overcome linguistic insecurity, the fear of making a mistake and showing oneself in public in a wrong practice. These considerations reinforce the interest of VLT as a complementary practice that, without replacing the classroom or immersion in language use environments (MONFERRER-PALMER, 2022, p. 77), contributes to a significant exercise on key aspects of use.

Although one might think that students prefer very short and specific videos, concrete and brief mini-pills have not been rated so positively by the surveyed VLT users. Another prejudice that has been contradicted is the assessment of the technical quality of the videotutorials. The vast majority of respondents prioritized the usefulness of the content of the videotutorial over its audiovisual quality.

In response to the question "What can be done to make the most of the videotutorial didactic tool in the context of language teaching?", it is a fact that didactic experts can learn from didactic self-taught people who, without studies in the subject, have become great teachers and problem solvers of all kinds thanks to free videotutorials on the Internet. This is increasingly important because, as we have seen, the vast majority of young people inform themselves through videotutorials (on YouTube, TikTok or Instagram), above consulting written sources of information (websites, libraries, instruction manuals). In short, the VLT is probably the instructional text most consulted by young people, the first source of information to resolve doubts of any kind, including linguistic ones.

On the other hand, it is dangerous for consumers of videotutorials to rely more on the communicative skills of the speaker than on the prestige or training in the professional field in question that the broadcaster has. This can be a problem, since the formal appearance of a videotutorial, its technical quality, can give the impression that its content is truthful, which could result in the dissemination of unreliable information. The information morass in which consumers are not prepared to consume content critically is the perfect breeding ground for the spread of *fake news* in the post-truth era (BUCKINHAM, 2019). It is also the role of the language teacher to teach their adult or adolescent learners to turn to reliable online sources of reference, including some language videotutorial accounts on social media.

In short, we can affirm that videotutorials in social networks are a hypergenre (MIRANDA, 2010) that summons other genres of self-regulated multimodal text in which there is interaction, although they are not based on conventional modes of interaction in the classroom. Videotutorials are a cistern from which to draw ideas and didactic techniques that can be useful for the classroom. It is necessary for teachers to know the VLT ecosystem on the Internet for different reasons: *a)* to update their knowledge of specialty, *b)* and didactic mechanisms, *c)* to know how to guide their students on useful and reliable videotutorial channels in relation to the competence of *learning to learn* or *knowing how to learn* that is included in the CEFR (COUNCIL OF EUROPE, 2002, p. 139); as well as in relation to the *can-do-statements* of the audiovisual competence described in the *Companion* (COUNCIL OF EUROPE, 2018, p. 68, 98).

In the first language videotutorials, which used to be published on the YouTube platform, and previously in videos or television programs such as That's English on Televisión Española, the set included the presence of a whiteboard in the background. The teacher, in an American shot and slightly displaced to one side, used it as a support for the explanations. This framing has been evolving towards post-production with graphics and towards the increasingly prominent weight of the vertical format, with a centered American shot of the teacher accompanied by graphics superimposed on the person that replace the whiteboard. Vertical videos are becoming a widespread format for the first time in the history of audiovisuals.

Unlike other types of videotutorials, VLTs do not usually have subtitles as such, but are accompanied by graphic elements to underline certain structures[3]. The most traditional form in which they appear is on a

---

[3] There are two types of VLT: those for learners with a specific source language and those for learners with any source language. In the latter case, subtitles can be important, but their use varies by platform. On YouTube subtitles are common, while on TikTok and Instagram, graphics are used instead.

whiteboard, although they are evolving towards a more sophisticated graphic post-production. In the case of the VLTs, there is a clear difference between those for beginners (A2-B1) and the advanced ones for preparing diplomas (C1-C2 in Catalan or Cambridge Certificates in English). In general, they all use a standard language model, although even more so in the VLTs for advanced levels. In any case, the normative correction is part of the essence of the vast majority of language videotutorials.

The platform conditions the formal characteristics of the VLTs: on YouTube they are horizontal and longer, while on Instagram and TikTok they are vertical and shorter. There is a contradiction that plays against the complexity of the communicative and linguistic phenomenon: there are platforms that increasingly demand shorter videos and more specific information, which makes it difficult to explain complex and nuanced concepts.

As for the conclusions referred to didactic modeling, we can affirm that the VLT has particularities over other videotutorials in terms of the area of work proposed, for the thematic content dealt with in relation to the use and knowledge of a language, for the characteristics of the metaverbal unfolding and for the use of a metalanguage to analyze the practices present by way of example. It is therefore a specific text hypergenre. The analysis of the reception of VLT shows the main interests of the users: complete reception of relatively long videos, preference for YouTube, choice according to their own obstacles and self-learning objectives, interactive involvement and interest in didactic potential beyond the technical audiovisual quality.

Finally, the survey of uses offers an orientation on how to improve the realization of the VLTs: constituting a didactic sequence of self-learning by the fact of the unfolding and the multiple possibilities to involve the receiver with the practice or with the desired learning. It should be made clear that this has been a preliminary study on the genre of videotutorial text for language teaching with the intention of formally and contextually defining the object of study from the ISD. It remains to elaborate a systematic and contrasted didactic modeling of the videotutorial based on second-generation didactic engineering and considering not only the passive use of videotutorials to reinforce content or to help students overcome learning obstacles in the context of attention to diversity, but also the elaboration of videotutorials by students as the final product of didactic sequences.

References

BUCKINHAM, David. Teaching media in a 'post-truth' age: fake news, media bias and the challenge for media/digital literacy education. *Culture and Education*, London, v. 31, p. 213-231, 2019.

CHARAUDEAU, Patrick. *Le discours d'information médiatique*: la construction du miroir social. Paris: Nathan, 1997.

COUNCIL OF EUROPE. *Common European framework of reference for languages: learning, teaching, assessment*. Strasbourg, FR: Council of Europe, 2018. CEFR companion volume with new descriptors. Available in: https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989. Access at: 11 Dec. 2022.

COUNCIL OF EUROPE. *Marco común europeo de referencia para las lenguas*: aprendizaje, enseñanza, evaluación. Traducción del Centro Virtual Cervantes. Strasbourg, FR: Language Policy División Strasbourg, Council of Europe, 2002. Available in: https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf. Access at: 11 Dec. 2022.

DOLZ, Joaquim. L'obstacle dans l'enseignement des langues: conceptualisation et opérationnalisation didactique. *En*: BALSLEV, Kristine; BRONCKART, Ecaterina B.; LAURENS, Véronique; NICOLAS, Laura (éd.). *Les obstacles dans l'enseignement des langues et dans la formation des enseignants*. Limoges: Éditions Lambert-Lucas, 2022. p. 300-310.

GROMIK, Nicolas. Video tutorials, Camtasia in the EFL classroom. *The JALT CALL Journal*, Tokio, v. 3, p. 132-140, 2007.

MIRANDA, Florencia. *Textos e géneros em diálogo*: uma abordagem linguística da intertextualização. Lisboa: Fundação Calouste Gulbenkian, 2010.

MOGOS, Andreea; TROFIN, Constantin. You Tube video genres: amateur how-to videos versus Professional tutorials. *Acta Universitatis Danubius*. Communicatio, Romania, v. 9, n. 2, p. 38-48, 2015.

MONFERRER-PALMER, Aina. *Docentes hiperdiscursivos*. Barcelona: Octaedro, 2022.

MOSTACERO, Rudy. Oralidad, escritura y escrituralidad. *Sapiens*. Revista Universitaria de Investigación, Chivilcoy, AR, v. 5, n. 1, p. 53-75, 2004.

OLIVEIRA, Lidineia; DIAS, Anair. Do gênero conto ao gênero podcast: uma proposta de retextualização. *Revista Educação em Foco*, Juiz de Fora, MG, v. 3, n. 26, p. 1-13, 2021.

OMIZO, Ryan Masaaki. *Facing vernacular video*. 2012. Tesis (Doctoral in Philosophy) - The Ohio State University, Columbus, OH, 2012. Available in: https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=osu1339184415&disposition=inline. Access at: 11 Dec. 2022.

PAIVIO, Allan. *Imagery and verbal processes*. New York: Holt, Rinehart and Winston, 1971.

ROMÁN, María Pilar. El vídeo tutorial como género discursivo en la proyección y desarrollo de la competencia oral en Secundaria. *Didacticae*, Barcelona, n. 8, p. 193-208, 2018.

TAPIA, Luis José; DELGADO, Nuria; GARCÍA, Leticia. Asimilación de contenidos y aprendizaje mediante el uso del video tutorial. *Latin American Journal of Science of Education*, Ciudad de Mexico, v. 7, n. 1, p. 1-7, mayo, 2020. Available in: https://www.lajse.org/may20/2020_12028.pdf. Access at: 11 Dec. 2022.

TARQUINI, Gianna; MCDORMAN, Richard. Video tutorials: an expanding audiovisual genre. *The Journal of Specialised Translation*, London, n. 32, p. 146-170, Jul. 2019. Available in: https://jostrans.org/issue32/art_tarquini.pdf. Access at: 11 Dec. 2022.

VYGOTSKI, Lev. *Pensée et langage*. Paris: La Dispute, 2019.

ZAYAS, Felipe; CAMPS, Anna (ed.). *Secuencias didácticas para aprender gramática*. Barcelona: Graó, 2006.