Multimodal Discourse Analysis Of Selected Videos For Teaching Vocabulary To Efl College Students In Iraq

Hanan Al-Radhi (PH.D)   University of Baghdad, Iraq halradhi2012@gmail.com

Abstract

A number of methods and techniques are explored for developing the vocabulary of ESL learners at all levels. Nowadays, YouTube videos are publicly available and students often watch the videos for learning various aspects of language including vocabulary. YouTube videos are not necessarily prepared by experts, and learners are not to decide their standard and quality. The videos use multi-modes like visual images, sounds, music, movement, instructions, and animation; hence, they are engaging, attractive, and interesting. All the modes contribute to the total effect of the videos. A multimodal discourse analysis helps to establish their standard and quality and suggests improvements to increase their benefits and utility for learners. This study examines a selected video and reviews its strengths and weaknesses.

Key words: YouTube, videos, discourse, multi-modes, analysis.

1. Introduction

The evolution in technology and COVID 19 pandemic have led to the creation of a bulk of learning resources for all subjects at all levels of education. Skill-based subjects are also being taught online. Teachers and educational experts have combined efforts to develop audio-visual aids that are intended for self-learning. Most of these are short videos that make good use of multiple modes like sound, images as well as instruction. The teacher or instructor directly addresses the learners and presents content or concepts as in a live classroom. While presenting, he/she receives help of images or sound tracks. Hence, the impact of the lesson is the total effect of all the modes, including the teacher's verbal, non-verbal instructions, the sound/music/songs/ and the images, photographs, and content in the form of Powerpoint. In order to understand the complete effect of a lesson, it is necessary to identify and isolate each kind of mode used during the lesson and to analyze it. Since the mode is multiple, the discourse analysis must also be multi-modal. This study is based on this assumption.

1.1. Background of the study

This is an age of multi-modes, whether someone thinks of trade and commerce, education, or social media. It has been observed that not only the youngsters are getting increasingly engaged in multiple modes, but also people of all age groups are attracted to them. Small children are highly impressed by the audio-visual modes compared to the conventional teaching-learning methods. Almost every individual is constantly occupied

with some electronic device. This obsession with devices can be fruitfully utilized for teaching purposes, and teachers are free to create and upload videos, free of charge, and make them open for public viewing. Especially during the lockdown period of COVID-19, when educational institutions were closed down, students could be benefitted by learning lessons from the self-learning videos. Online sessions have problems like connectivity, some degree of knowledge of computers, and time slots. However, the videos are available anywhere and at any time. There is no harm if teachers check the quality and standard of the videos and recommend them to their students. The quality and standard have to be established and verified, which can be achieved through discourse analysis (DA hereinafter).

1.2. Importance of the topic

The current age is an age of multi-modal channels. Digital devices are buzzing with videos of all kinds. Videos can be effectively used for educational purposes to teach both content-based subjects as well as skill-based subjects like languages. Multi-modal content is being prepared for developing all four language skills of learners, particularly, second language learning, which has always been a challenge for teachers and instructors. Multi-modes include a voice track or direct address by teachers/instructors, images, and sound. Hence, the impact of a lesson created in multi-modes depends upon the combined effect of all the modes. Besides verbal and non-verbal interaction, there are audio elements like sounds and music or visual elements like pictures, images, photographs, charts, slides or videos, or even occasionally direct filming of situations. The attempt is to make learning simple and quick by using each mode optimally. When written content is used for instruction or a lesson is delivered by a teacher, the content is subjected to discourse analysis to investigate its strength and weaknesses. Similarly, it is necessary to conduct discourse analysis of multi-modal content, which is the new order.

Multi-modal content analysis will lead to an understanding of the impact of each mode separately and as a unit. It will allow content-makers to modify the content. Standard multi-modal content can be safely given to learners. The quality and standard can be established through DA. Also, which mode is suitable for which kind of language skill will be revealed. Hence, the content makers will be able to decide how much content of which mode should be used for a given lesson. Taking the current circumstances into consideration, it can be said that the future of content is multi-modal.

1.3. Hypothesis

Multimodal discourse analysis of videos helps to assess and enhance the quality and effectiveness of videos used for teaching vocabulary.

1.4. Research questions

1.4.1. How are videos useful for teaching vocabulary in EFL primary classes in Iraq?

1.4.2. How can multimodal discourse analysis of videos help to assess their effectiveness in teaching vocabulary?

1.4.3. In what way can multimodal discourse analysis lead to improve and enhance the quality and effectiveness of videos?

2. Theoretical Background

2.1. Meaning and nature of multi-modal content

"Content" here, refers to the content that is prepared or created for the purpose of teaching or instruction. Traditional content contained a single mode, i.e. the written mode. There is general familiarity with text-books used for teaching. The text books contain written content. Reviewing text-books and analysing content is a favourite concern of researchers. However, in recent times, the form of content has changed. Printed and written content is being replaced by multi-modal content to which learners respond virtually.

Cope and Kalantzis, (2009) state that multi-modal content is a combination of two or more modes like spoken and written language, fixed images, moving images, audio tracks, gestures and spatial meaning. Multi-modal content transmits a message through multiple modes. Since the use of multi-modal creation is increasing rapidly, it is necessary to review and analyze the content in order to avoid danger of sub-standard or incorrect content placed in the hands of learners. Every individual has free access to platforms on which he or she can prepare and upload educational videos that are open to the public. These are neither censored nor reviewed, and each teacher is suggested to be the best. Due to teachers' enthusiasm to impart knowledge, they prepare videos and upload them, sometimes just to gain popularity regardless of any genuine concern for education. Hence, the quality and standard of these videos needs scrutiny.

Multi-modal content includes a variety of modes, both auditory and visual. They can be made very attractive by exploiting technology, and along with images and slides, it is possible to add background music, live scenes, light effects, and sounds. Young learners often get engaged in these effects and the core lesson gets neglected. Thus, the excessive use of technology is not advisable while preparing multi-modal content – it must be used discretely, responsibly, and sparsely, so that it does not destroy the real purpose of the content.

Multi-modal content has a number of advantages. The same content is handled in different ways in several modes. Thus, each learner can follow the mode they find suitable. "Learners have a preferred learning modality such as visual, auditory, or kinaesthetic, while many learners are multimodal (use a combination of these modalities)" (Gilakjani et al, 2011). All senses are used by learners to react to the content thus keeping them engaged and involved. There is a greater interaction with the content because learners respond to multiple modes. Such an interaction provides a scope for flexibility and creativity for teachers / instructors.

2.2. Preparing multi-modal content

Since educationists are new to the creation of multi-modal content which has gained popularity recently, it is necessary to deal with the task with a sense of responsibility and carefulness. Educationists must be aware of and have knowledge about all the modes available for teaching. This will ensure that they combine the modes judiciously. It is not necessary to use every mode in each lesson – only suitable and appropriate modes depending on the topic must be selected. Care should be taken as videos should not become very loud and too attractive as there is a chance of learners being carried away by the effects while the actual lesson is ignored.

Visual content gives scope for size, colour, focus and texture. "We are entering a historical epoch in which the image will take over from the written word (Gombrich, 1996)."

The background can also be effective. Videos offer control over speed, position of the camera, angle, and perspective. In audio tracks, there is scope for volume, speed, music, tone, and frequency. The final stage is editing the videos. All the content in the videos must be logically arranged. Videos must be checked for the correctness of content, limits of the content, appropriateness as per age-group and topic and combined effect of all modes.

2.3. Teaching Vocabulary

Educational achievements rely on the size of vocabulary mastered by a learner. This statement is applied not only on language skills including both the oral and the written, but to all subjects including science and social studies. The vocabulary necessary for routine activities is often picked up incidentally by second language learners; however, technical, abstract, and complex vocabulary has to be taught explicitly and deliberately. Some vocabulary is basic, which includes words that are used in everyday life. Some words are necessary to express opinions, to comment, or argue. The third tier of vocabulary includes a reservation of words used for specific purposes, like academic terms and technical words.

While teaching vocabulary, the following aspects need to be covered: meaning, pronunciation, spelling, context (register) and usage. Spelling is also important in this digital age where computers often tend to autocorrect spellings and convey a wrong meaning. Words may also have different connotations and denotations.

Teaching words without a context and for the sake of teaching them only becomes a meaningless activity in a classroom. Hence, it is important to make vocabulary learning interesting by embedding the words in context, conveying their usage through stories, jokes etc., and giving the students an opportunity to use the words. Various activities can be conducted for teaching vocabulary. Games are used for teaching vocabulary. Rich vocabulary is a basic necessity for competent use of language. "Without grammar, very little can be conveyed, and without vocabulary, nothing can be conveyed" (Thorbury, 2002)."Vocabulary supports the mastering of the four language skills" (Rahmadhani, 2015).

In recent times, a bunch of videos have been uploaded through which learners can develop vocabulary independently. The young generation seems to have a preference for being active on digital devices and find it comfortable to learn language through videos.

2.3.1 Learning vocabulary through multi-media

YouTube videos are a good example of multi-media for teaching language skills including grammar and vocabulary. The videos make use of graphics, visuals, actions, auditory, and direct instruction. All modes can be combined. The videos are short – generally of ten to fifteen minutes duration. By this, they ensure the attention span of learners at all levels. Only a limited number of words can be covered in the short period which prevents the overload of information. Multi-modal videos have a very good potential as a method for teaching vocabulary to second language learners. Hence, they can be utilized for having the best results. According to Juhana (2014), learning vocabulary can be made interesting and enjoyable through YouTube videos. Riswandi (2016) defined YouTube as a website that allows users to upload and watch videos. They can also share the videos they like with other people. Also, using videos in the teaching- learning process does assist students because these videos are able to shed light on all fields of knowledge and interests that might grab students' attention (Kabooha & Elyas, 2018).

While selecting and using videos for vocabulary learning, it is necessary to verify their standard and quality, correctness, appropriateness and total effect. They must also be attentive when it comes to the evaluation of vocabulary in order to be considered as a comprehensive means of vocabulary learning. They encourage self-learning and independent learning. Not only are there videos for learners to help them acquire vocabulary, but also there are videos for teachers and instructors that tell them various ideas about teaching vocabulary, so we are concerned with the self-learning videos in this study.

2.3.2 YouTube videos

YouTube is a popular social platform for people of all age groups. All kinds of content related to any topic can be published. Also, there are worldwide viewers; hence, it is not a concern whether the videos would find viewers or not. Today, there are billions of YouTube viewers. This fact increases business possibilities, and the business earnings depend upon the popularity, so many people have made it their full-time profession. There is a provision of interaction through the comment section. Users can click on the like or dislike buttons, subscribe to channels and write comments.

Of course, there is some risk involved. The videos can be good or bad, anonymous people can become celebrities overnight. Despite the fact that YouTube has the age restrictions feature, the youth are often driven away by trending videos which might not be appropriate form them. But still, YouTube videos can be used as much for both: instructional and entertaining purposes.

3. Literature Review

Gilkajani (2011) and many other specialists studied the 'Effects of multimodal learning models on language learning and teaching'. They pointed out that the arrival of computers has changed our perspectives regarding language and communication. The change has also affected language learning. These specialists have discussed multimodality in relation to second language acquisition in the article. According to them, following are the principles of multimedia:

- Words supported by pictures are better than mere words.
- Learners' attention is focussed in multimedia learning; hence, it is not split.
- Multimedia content should exclude irrelevant information.
- Interactive media in the learning process is more efficient.
- Learner knowledge structures should be activated before the learners are exposed to multi-modal content for better results.
- Animated multimedia improves learning.
- Engagement of the learners in the content must be ensured.
- Feedback and opportunity to apply what acquired proves to be effective.

The article discusses in details how the computer has assisted language learning and how it has boosted its advantages. The meaning and implications of multimedia learning have also been discussed. Learning is not merely a visual-cognitive activity; it includes the interplay of all sensory organs. Furthermore, learning is a kind of life experience wherein we use all senses and not just vision, which was the main concern in traditional education. Language, or second language, is also a life experience. Multimedia illustrations can be a great means to support the teaching- learning environment of the second language acquisition platforms. Having an interaction between the teacher of the second language and the student via pictures, stories, or even animations can help ensure the elements of the communication skills. The writers conclude that multi-media content compensates the absence of certain cognitive structures in the learners; thus, enabling them to understand better according to their own capacities. "Attention to the meaning-making potential of the various multimodal designs can help language learners to cope more efficiently as they face new modes of information presentation "(p 1326).

Magnusson and Godhe (2019) have written an article titled 'Multimodality in Language Education – Implications for Teaching'. They claim that due to digitalization, the content of language education has been challenged. They viewed two poems from multimodal perspectives. On the basis of their findings, they tried to understand the challenges in teaching. They concluded that by stressing the importance of multimodal content in meaning-making and language teaching. The very nature of communication is affected because of the way the youngsters interact with a combination of modes, especially in informal situations. The study is based on multimodal socio-semiotic theory. It is observed that in spite of multi modes and digital platforms, the traditional forms of verbal reading and writing still retain their importance, especially in poetry teaching. Multimodal theory was developed in the1990s and is currently undergoing a number of changes. It is being demonstrated for its possibilities and limitations. The shift between

modes is possible through digitalization. Full meaning of content can be conveyed through all modes, that is textual meaning, ideational meaning and interpersonal meaning. With multi-modes, learning can go beyond knowledge transfer. The writers point out that the process of learning cannot be observed, but the change brought about as a result of learning can. Through the study of two poems from multimodal resources, the research shows how to assemble various modes. The research was conducted in a Swedish school. In conclusion, the writers state that in this era, it is necessary to understand and use multi-modes for meaning -making and language teaching.

Hariyono (2020) deals with teaching English vocabulary to pupils using YouTube videos. He conducted a descriptive-qualitative research, and the data were based on observation. The study was conducted at Bogor, and the sample consisted of seven students studying in grade two. The writer points out that there is easy and quick access to YouTube videos, and they make language learning fun. Videos change the learning environment. The treatment consisted of showing YouTube videos to the students. Through observation, the researcher noted the achievements of the students in vocabulary. The students' response was positive, they remained engaged in the videos, they asked questions eagerly, and they were excited about the videos. They enjoyed using technology for study. Therefore, the task of learning vocabulary proved meaningful due to the videos. The researcher concluded that YouTube videos are one way to improve the vocabulary of young students studying English as a second language.

"Discourse and Technology – Multimodal Discourse Analysis" (2004) is a book edited by Philip Le Vine and Ron Scollon. It includes chosen papers from the Georgetown Roundtable on the field of Linguistics held in 2002. In the first article that serves as an introduction to the book, the editors refer to Multimodal discourse analysis (MDA hereinafter) as a confluence of discourse and technology. They consider MDA as a result of technological change. They emphasize the fact that language always functions across multiple modes. They point out the importance of studying discourse and technology along with MDA. The assumption of all the authors who have contributed to the book is that 'discourse is essentially multimodal, not monomodal (p 11).' Social actions are also looked upon as multimodal phenomena. Theu Van Leeuwen has given ten reasons in his article why linguists should shed light on communication especially the visual one. He gives the example of a poster that contains a speech act and at the same time an image act. The message that is communicated is an impact of both. He concludes that multimodal analysis is a new and exciting area for research in which many treasures can be uncovered. Rodney Jones, in his article, has pointed out the problem of context involved in computer use for communication. Previously, the tendency was to focus on written and spoken texts while objects and actions were considered to be part of the environment (Halliday, 1989). It is difficult to decide which aspects of a situation can be considered as content and even harder for a researcher to get ingress to the different communication features. Linguistic analysts need to change their perspectives, include a variety of modes, and make good use of different approaches. Saint Georges' article is about the effect of space and layout in producing meaning'. Discourse takes place in social

and physical space. And Space affects interaction. According to the writer, DA has not traditionally paid much attention to the role of space. Participants move across spaces and perform actions. This affects the impact of interaction. Spatial layout also affects discourse. The writer quotes Kress and Van Leuwen (p 83) - "Multimodal approaches to discourse point to the fact that utterances are only a moment in the continuous process of communication and that there is no necessary priority of language over other modes of meaning making in social actions (Kress et al. 2001; Kress and Van Leuwen 2001)." The conclusion is that discourse has to define an event as taking place in space. Discourse cannot analyze interactions solitarily and away from physical spaces. Laurent Filliettaz has investigated the 'Multimodal Negotiation of Service Encounters'. His focus is on non-verbal behaviour during negotiations.

An article that is more significant for this study is Sigrid Norris's' "Multimodal Discourse Analysis: A Conceptual Framework". Traditionally, discourse revolves around spoken language. The data in this study were collected with a video camera from natural interactions. The framework was extended to include 'all identifiable communicative modes'. The modes cannot be easily separated. Kress (2001) and many other specialists consider that communication is attained through different modes whether these modes are put together or handled seperately. The writer takes a heuristic view of all communicative modes while analysing situations. Elaine Yakura has also considered this in "Informed Consent" and "Other Ethical Conundrums in Videotaping Interactions."

Carey Jewitt has written about "Multimodality and New Communication Technologies". He talks about medium and mode. While medium is related to how selections are spread, such as printed books, digitalized apps, mode is related to any assembled, normal way of illustration and communication, such as pictures, movements, pieces of music or writing, or new arrangements of the elements of these (Kress, 2001). In this age, we observe a movement from text to screen. He concludes that meaning can be designed in different ways through multiple modes. Frederick Erikson's article is titled "Origins: A Brief Intellectual and Technological History of the Emergence of Multimodal Discourse Analysis." In his article, he takes a historical survey of DA and states that audio-visual content requires multimodal analysis. On the whole, the book consists of articles that tackle most aspects of multi-modal discourse and its analysis.

4. Methodology

4.1. The Method of MDA

MDA studies how different modes like pictures, images, videos, and sounds combine with words to make meaning. According to Halliday's social semiotic approach to language, people use different semiotic resources like music, gestures, and sounds to create meaning. According to Halliday, language serves three types of functions:

4.1.1. Ideational (the topic of the content).
4.1.2. Interpersonal (relationship between the participants).
4.1.3. Textual meaning (how the text is organized).

Multimodal discourse involves print genres as well as audio-visual genres. According to Van Leeuwen, there is a relationship between speech acts and genre with relation to multimodality. It explains the 'how' versus the 'what' of multimodal communication. It may be an illocutionary act or perlocutionary act. Understanding an element of vocabulary is an illocutionary act and using it is a perlocutionary act.

Communication modes as data types for analysis include:

Process:
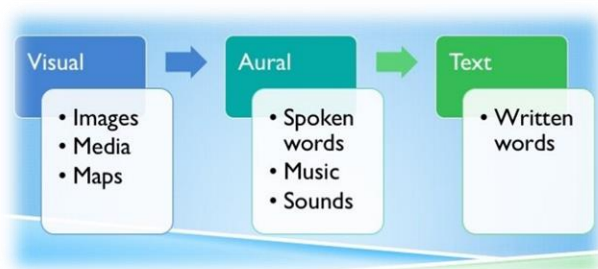
1. Identifying the modes employed.
2. Explaining the function of each mode.
3. Concluding the meaning derived.

4.2. Limitations of MDA:

4.2.1. It is time consuming.
4.2.2. It is often very technical.
4.2.3. Language aspects tend to be neglected.
4.2.4. Wider social issues may be involved.
4.3. Approaches to MDA

4.3.1. Systemic functional multimodal discourse analysis (SF-MDA)

SFL is designed by Michael Halliday who perceived language as a social system, a meaning making process. SFL has developed a systemic functional grammar (SFG) to illustrate the meaning-making potential of language and to provide a sample that helps to understand discourse. These dimensions are used in systemic functional multimodal discourse analysis.

4.3.2. Social semiotic multimodality

It deals with the processes involved in meaning-making. Individuals and societies shape language in order to realize their ideologies. It examines how modes serve social needs

in a context. It relates to systematic functional linguistics (SFL) and critical analysis discourse (CDA).

4.3.3. Multimodal interactional analysis.

In early times, the focus was on the analysis of conversations, but now it has broadened its scope to include videos consisting of multimodal interactions. The focus is on coordination and sequence of action, gestures, eye-contacts, movement, posture, use of space and objects.

5.   Principles for analysis and the sample

This study is based on the principles presented in the book of Biber and Conrad (2009), Register, Genre and Style. The type of discourse used by YouTubers can be characterized as a unique register itself. Biber and Conrad (2009) define register as a diversity accompanied with a specific condition of use, encompassing specific communicative aims.

According to Biber, the first principle is 'situational'. It is about the corporeal context, the genuine setting, register, and genre. The study will use the following pattern proposed by Biber and Conrad (2009): participants, connections among participants, mode, creating conditions, time and place, communicative goal and point.

And when it comes to the sample, the following video is selected for analysis.

https://youtu.be/h_i2jDbMGwA

5.   Data Analysis

Content: This video teaches ten words. Although the title of the video tells us that it is about correct pronunciation of words, the video gives the meaning, contains examples of usage of the words in sentences and additional information that is useful in fixing the meaning of the word, as well as differentiating it from words with which it can be confused. The inclusion of comprehensive content is the reason why this video has been selected in this research.

The following words are included in the video:

Silicone, silicon, mischievous, etcetera, nuptials, pronunciation, parenthesis/ parentheses, strengths, jewellery, anyway.

5.1.   Participants

The participants are scattered worldwide. The teacher or the instructor speaks in the video. She mentions that her student is an actress. The actress is also one of the participants. The actress is called Katura, and she is asked to pronounce the words.

5.2.   Interaction

The flow of conversation is one to one, yet the instructor takes for granted a participative audience, thus adding to her zest for teaching. She builds a rapport from the beginning by declaring that she has selected words that are often mispronounced even by native speakers or advanced users of English language. This strategy assures that viewers would continue to watch the video. She warns listeners against imitating someone without being careful and suggests that correct pronunciation of the words would make the viewers sound professional and educated.

These two traits work easily because who would not want to sound professional and educated? Hence, although the flow is in one direction, she can ensure active participation and listening from the viewers thus establishing an indirect interaction.

There is a direct interaction between the instructor and the actress. There is self-reference several times throughout the video. The word 'I' points towards the instructor in the beginning and also toward the end. In the beginning, she mentions a number of times that the viewers must watch the video till the end. She says that she had a surprise sentence at the end that would tell them more about pronunciation. This is another tactic of making sure that the entire video is watched and also confirming audience for the coming future videos. She appeals to the viewers to subscribe to her channel.

The viewers cannot participate in the video, so their feedback would be late coming from the comment box. Hence, she asks Katura about her experience. Katura reacts positively assuring her that it was fun, more like a game than like a lesson. Katura's feedback helps to condition the minds of the viewers to some extent.

5.3. Mode

It is a spoken mode. The instructor utters the words, their meaning, and sentences containing the words as examples, and some additional information about them. She reads out the words and sentences that appear in print along. It is more or less like a classroom, and the background where the words / sentences r serve the same purpose as the whiteboard in the classroom. However, it is not a face-to-face communication but communication through electronic media.

5.4. Production circumstances

Natural speech is spontaneous and unplanned. There may be repetitions or corrections. In the video, the speech is filmed; hence, it is consciously planned. The speed and utterances of the teacher match with the display of images, printed texts on the screen.

5.5. Setting

"The setting refers to the physical context of the communication – the time and place. An important consideration here is whether the time and place are shared" (Biber 2009). The setting might be a studio or simply any room where the video is filmed. The setting for each sentence is seen through the images, and it is different when Katura speaks. In the video, the time and place are not shared. The video is made in accordance with the convenience of the teacher, and it can be watched anywhere and at any time by anyone.

5.6. Topic

The topic is an introduction to ten selected words, their meaning as particularly, and their pronunciation.

5.7. Purpose of communication

The purpose is academic. The teacher tries to teach the new words and their pronunciation to the students so that they should not mispronounce the words and as part of the vocabulary development process.

5.8. Multimodal Features

This section analyses the elements of multimode that are used in the video. It also includes elements of pragmatics and discourse analysis.

5.8.1. Visual features

The main screen, or the title screen focuses on the word 'pronunciation' and the image of the instructor. The image of a conference table with a green plant in a corner is displayed continuously unless some other image is deliberately shown. The table image on the right and the instructor on the left, is the position maintained throughout the video. All typed sentences are opposite to the background of the image, or opposite to the special images shown as per the content.

When the actress, Katura, is introduced, her photographs are shown. Almost every utterance is accompanied by an appropriate image like a revolver when referring to a criminal, a company when referring to a businessman, or a car when referring to a mechanic. The images for the words chosen are also apt. Additional information about the words is given through the images like the photograph of Silicon Valley in California. After the display of the meaning of each word, the sentence given as an example is typed on the screen. The explanation of each word ends with the phonemic transcription of the word displayed on the screen. All the images / photographs are fixed.

5.8.2. Audio- features:

There is no music in the video. The voice of the instructor and the voice of the actress are the only sounds. The actress laughs heartily every time she mispronounces a word or while giving her feedback. The speech of the instructor is clear. The pronunciation is repeated several times. She maintains a steady speed throughout – not too fast and not very slow. When the instructor asks the viewers to click the bell icon, the sound of a bell is heard. These are the only sound effects in the video. Which gives the video a formal touch and academic aspect.

5.8.3. Pragmatic features

If the students understand the meaning as well as use the word, it can be said that the video fulfils the illocutionary and perlocutionary speech acts. The instructor is polite. She gives enough time for the actress to take her turn in speaking. The overall tone of the speaker is polite; hence, in spite of her pointing out incorrect ways in which people

pronounce words, it does not annoy or sound rude. Since the very aim is to explain the meaning of words, there are no implied meanings. She means what she says.

5.8.4. Non-verbal features

The instructor uses gestures, and actions to complement her speech. Her facial expressions convey her message clearly. The eyes look directly at the viewer – she is positioned in the front and at a chosen height. The gestures appear to be natural, but they may be practiced. There is no movement in the given space.

5.9. Discussion:

The background images help fixation and better understanding of the concept. Images must be displayed in accordance with the sentence content, which we find in this video.

The actual use of the new word in various contexts is essential and reveals the usage of the word in sentences. Examples are the best way of introducing vocabulary. Sufficient examples are included to explain the meaning of each word. All aspects of vocabulary – spelling, pronunciation, meaning, and usage are taken care of. The last part draws attention towards another important aspect; and this is word stress. This motivates the viewers to await the next video.

At some places, the typed text covers the images partly. This is bound to happen because the screen is small. However, care is taken, and the image is displayed for enough time on the screen before the text appears.

The words follow each other in succession, but there is enough time to allow the viewers to ponder over each word. A little bit of animation could have added to the total effect and made the videos more attractive. However, there must not be much animation because this might distract the attention of learners. Similarly, some sound effects could be added discretely. Maybe, continuous soft music in the background and a popping sound every time a new word is introduced are multi-modal features that might have scope in YouTube Videos.

The examples were apt and correct. There were no spelling or grammatical errors. The duration of the video is ten minutes for ten words which suits the attention span of college level students.

6. Conclusion

From the multi-modal discourse analysis of the selected YouTube video for self-learning of vocabulary for college students, it can be concluded that YouTube videos are a good way for teaching English vocabulary at the college level. They must be prepared by experts, not only language experts, but expert technicians so that more attractive animations and images can be used.

It is advisable to include some form of direct interaction. Beginning with a greeting and an introduction to the topic is necessary. Interaction and self-reference are likely to

fetch more comments because they add a personal touch to the videos. The teacher can also suggest some home assignments. Also, a reference to the next video can ensure participation so that students will learn more vocabulary.

The study shows that a multi-modal discourse analysis of a YouTube Video helps to note the strengths and weaknesses of each mode separately with reference to the teaching unit. In addition, the total effect of the combination of all modes can be understood. The analysis helps to suggest changes and modification, additions and deletions in order to make the videos more effective, beneficial and fruitful for learners.

References

Biber, Douglas et al. Register, Genre, And Style.

Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New Literacies, New Learning. Pedagogies: An International Journal, 4(3), 164–195. https://doi.org/10.1080/15544800903076044

Filliettaz, L. (2009). Les discours de consignes en formation professionnelle initiale : une approche linguistique et interactionnelle. Éducation et didactique, (3-1), 91–120. https://doi.org/10.4000/educationdidactique.445

Gilakjani, Abbas Pourhossein et al. "The Effect Of Multimodal Learning Models On Language Teaching And Learning". Theory And Practice In Language Studies, vol 1, no. 10, 2011. Academy Publication, https://doi.org/10.4304/tpls.1.10.1321-1327.

Gombrich, E. H, and Richard Woodfield. The Essential Gombrich. Phaidon Press, 1996.

Halliday, M. A. K. Spoken And Written Language. [World Publishing Corporation], 2012.

Holliday, M. (1989). Halliday on Spoken and Written Language. Reclaiming The Language for Social Justice. https://reclaimingthelanguage.wordpress.com/2018/09/01/halliday-on-spoken-and-written-language/

Hariyono, Tira Chairiah. "TEACHING VOCABULARY TO YOUNG LEARNER USING VIDEO ON YOUTUBE AT ENGLISH COURSE". Language Research Society, vol 1, no. 1, 2020. Universitas Presiden, https://doi.org/10.33021/lrs.v1i1.1038.

Jewitt, Carey. "The Move From Page To Screen: The Multimodal Reshaping Of School English". Visual Communication, vol 1, no. 2, 2002, pp. 171-195. SAGE Publications, https://doi.org/10.1177/147035720200100203.

Juhana. "Teaching English To Young Learners: Some Points To Be Considered". Asian Journal Of Education And E-Learning, vol 2, no. 1, 2014, pp. 43-46., Accessed 24 Jan 2022.

Kabooha, R., & Elyas, T. (2018). The Effects of YouTube in Multimedia Instruction for Vocabulary Learning: Perceptions of EFL Students and Teachers. English Language Teaching, 11(2), 72. https://doi.org/10.5539/elt.v11n2p72

Kumar, E. S. (2011). Computer Assisted Language Learning (CALL) (Chapter 1) - A Handbook for English Language Laboratories. Cambridge Core. https://www.cambridge.org/core/books/abs/handbook-for-english-language-laboratories/computer-assisted-language-learning-call/97FCB7A84A31665A92E7C3D2051D0D54

Kress, G., and C. Jewitt. "Computer Mediated Learning: The Multimodal Construction Of Mathematical Entities On Screen". 2022, Accessed 24 Jan 2022.

Kress, Gunther R. Multimodal Teaching And Learning. Bloomsbury Publishing PLC, 2014.

Kress, Gunther, and Theo van van Leeuwen. Reading Images. Routledge, 2021.

Lang, Peter. Multimodal Learning: Moving Beyond Language. 2022.

Magnusson, Petra, and Anna-Lena Godhe. "Multimodality In Language Education – Implications For Teaching". Designs For Learning, vol 11, no. 1, 2019, pp. 127-137. Stockholm University Press, https://doi.org/10.16993/dfl.127.

Magnusson, P., & Godhe, A.-L. (2019). Multimodality in Language Education – Implications for Teaching. Designs for Learning, 11(1), 127–137. https://doi.org/10.16993/dfl.127

Rehmadhani, P. A. "Techniques In Teaching Vocabulary To Young Learners At LIA English Course". TELL-US Journal, vol 1, no. 2, 2014, pp. 1-8., Accessed 24 Jan 2022.

Riswandi, D. "Use Of Youtube-Based Videos To Improve Students' Speaking Skill". The 2Nd International Conference On Teacher Training And Education, 2016, p. 269, Accessed 24 Jan 2022.

Syafrizal, Syafrizal, and Haerudin Haerudin. "THE IMPLEMENTATION OF VOCABULARY BUILDING STRATEGY IN TEACHING ENGLISH VOCABULARY TO YOUNG LEARNERS". Jo-ELT (Journal Of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP, vol 5, no. 1, 2018, p. 40. LPPM IKIP Mataram, https://doi.org/10.33394/jo-elt.v5i1.2296.

Thorbury. (2002). OSF. https://osf.io/p4xc9/download/?format=pdf#:~:text=Wilkins%20in%20Thornbury%20(2004:%2013,do%20not%20know%20many%20vocabularies.

Van Leeuwen, T. (2016). A Social Semiotic Theory of Synesthesia? - A Discussion Paper. HERMES - Journal of Language and Communication in Business, (55), 105. https://doi.org/10.7146/hjlcb.v0i55.24292