MULTIMODAL DISCOURSE ANALYSIS OF SELECTED VIDEOS FOR TEACHING VOCABULARY TO EFL PRIMARY STUDENTS IN IRAQ

Prof. Dr. Bushra Saadoon Mohammed Alnoori

Al-Masour University College, Iraq

DOI: http://doi.org/10.37648/ijrssh.v12i02.042

ABSTRACT

The growth in technology and COVID 19 pandemic has led to the creation of a bulk of learning resources for all subjects at all levels of education. Skills-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 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 takes help of images or sound tracks. Hence, the impact of the lesson is the total effects 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 ppt. 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 analyse it. Since the mode is multiple, the discourse analysis must also be multi-modal. This study is based on this assumption.

BACKGROUND OF THE STUDY

This is an age of multi-modes, whether one thinks of trade and commerce, education or social media. It has been observed that not only the youngsters but people of all age groups are getting increasingly engaged in multiple modes. 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 purpose and teachers have a free hand 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 are closed down, students can be benefitted by learning lessons from the self-learning videos. Online school 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.

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 together. 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 discourse analysis. Which mode is suitable for which kind of language skill will also be revealed. Hence, the content makers will be able to decide how much of content of which mode should be used for a given lesson. Judging by the present conditions, it can be said that the future of content is multi-modal.

Hypothesis

Multimodal discourse analysis of videos helps to assess and enhance the quality and

effectiveness of videos used for teaching vocabulary.

Research questions

1) How and to what extent are videos useful for teaching vocabulary in EFL primary classes in Iraq?

2) how can multimodal discourse analysis of videos help to assess their effectiveness in teaching vocabular?

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

THEORETICAL BACKGROUND

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, that is 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 the 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, still images, moving images, audio tracks, gestures and spatial meaning. Multimodal content communicates a message through multiple modes. Since the use of multi-modal creation and use is increasing rapidly, it is necessary to review and analyse the content lest there remains no danger of sub-standard or incorrect content going 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 under the notion mostly that he or she is the best. In their enthusiasm to impart knowledge, teachers prepare videos and upload them, sometimes more so to gain popularity than with any genuine concern for education. Hence, the quality and standard of these videos needs scrutiny.

Multi-modal content includes a variety of modes, both audio 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, sounds and so on. Young learners often get engaged in these effects and the core lesson gets neglected. 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. Hence, each learner can follow the mode that he finds suitable. "Learners have a preferred learning modality, namely, visual, aural, read/write or kinaesthetic, while many learners are multimodal (use a combination of these modalities)" (Gilakjani et al, 2011, p 1322). All senses are used by learners to react to the content thus keeping them engaged and involved. There is greater interaction with the content because learners respond to multiple modes. It provides scope for flexibility and creativity on the part of teachers / instructors.

Preparing multi-modal content

Since educationists are new to the creation of multi-modal content which has gained popularity recently, it is necessary to carry out the task with a sense of responsibility and carefulness. Educationists must be aware of, and have knowledge about all the modes available for the purpose of 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 that the videos do not become very loud and too attractive as

638

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. The background can also be made 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 correctness of content, limits of the content, appropriateness as per age-group and topic and combined effect of all modes.

Teaching Vocabulary

Educational achievements depend to a large extent on the size of vocabulary mastered by a learner. This statement applies not only to language skills like listening, speaking, reading and writing, but to all subjects including science and social studies. 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 in isolation for the sake of teaching them 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, without vocabulary, nothing can be conveyed" (Thorbury, 2002, p. 13)." Vocabulary supports the mastering of the four language skills (Rahmadhani, 2015).

In recent times, a host of videos are uploaded through which learners can develop vocabulary on their own, independently. The young generation seems to have a preference for being active

on digital devices and find it comfortable to learn language through videos.

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, audio and direct instruction. All modes can be combined. The videos are short – generally of ten to fifteen minutes duration. This they ensure attention span of learners at all levels. Only a limited number of words can be covered in the short period, this preventing 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 best results. According to Juhana (2014) learning vocabulary can be made interesting and enjoyable through YouTube videos. Riswandi (2016) defined YouTube as a video-sharing website in which users can upload, share and view videos. Using video in the learning process helps students because it

can explain all fields of knowledge and interest and the attention of students (Kabooha & Elyas, 2015).

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 take care of evaluation of vocabulary in order to be considered as comprehensive means of vocabulary learning. They encourage self-learning and independent learning. Just as there are videos for learners that help them to acquire vocabulary, there are videos for teachers and instructors that tell them various ideas for teaching vocabulary, we are concerned with the self-learning videos in this study.

YouTube videos

YouTube as a popular social platform for people of all age groups.

All kinds of content can be published related to any topic under the sun. there are worldwide viewers, hence, it is not a concern whether the videos will find viewers or not. today, there are billions of YouTube viewers and it has the advantage of business possibility. The earnings depend upon the popularity and there are people who 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 risk involved. The videos can be good or bad, anonymous

people can become celebrities overnight. The videos are not censored or marked for age-groups. Children cannot decide what is good or bad for them. Children and youth are often driven away by trending videos. YouTube videos can be used as much for instructional purpose as for entertainment.

Literature Review

Gilkajani et, (2011) studied the 'Effects of multimodal learning models on language learning and teaching. They pointed out that the arrival of computers has changed our perspectives about language and communication. The change has affected language learning. They 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; it is not split.
- Multimedia content should not include extraneous and redundant information.
- Interactive multimedia learning is more effective.
- 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 detail CALL (computer assisted language learning) and its advantages. The meaning and implications of multimedia learning are also discussed. Learning is not merely a visual-cognitive activity; it includes the interplay of all sensory organs. Learning is a kind of life experience wherein we use all senses and not just vision, which was the sole media in traditional education. "Language, or second language is also a life experience. Multimedia presentations can be an excellent means toward re-constructing a pseudo-natural environment in which these negotiations of meaning that serve as the platform for second language acquisition can take place. Accompanying the interaction between teacher and second language learners with a combination of modes such as animation (visuals and movement) and narration can provide a variety of contexts to be used for developing all the components of communicative competence (p 1325). The writers conclude that multi-media content compensates for 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 (1326)."

Magnusson P. and Godhe L (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 conclude 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 1990s and is currently undergoing a number of changes, is being explored 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 be observed. 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 age, it is necessary to understand and use multi-modes for meaning -making and language teaching.

Hariyono C. (2020) deals with the teaching of English vocabulary to young learners using YouTube videos. He conducted descriptive-qualitative research and the data was based on observation. The study was conducted at Bogor and the sample consisted of seven students studying in class 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. The task of learning vocabulary proved meaningful due to the videos. The researcher concluded that YouTube videos are one way of improving 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 selected papers from the Georgetown Roundtable on Languages and Linguistics held in 2002. In the first article that serves as an introduction to the book, the editors refer to Multimodal discourse analysis (MDA) as a confluence of discourse and technology. MDA is 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 and 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 pay attention to visual communication. 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 area, an 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 mediated 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, 1975, p 25). It is difficult to decide which aspects of a situation can be considered as content and even more difficult for an analyst to get access to the various aspects of communication. Linguistic analysts need to change their perspectives, encompass multiple modes, and make use of multiple methods. Saint Georges' article is about 'the influence of space and layout in making meaning. Discourse takes place in social and physical space. 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 independent of physical spaces. Laurent Filliettaz has investigated the 'Multilodal 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 centred around spoken language. The data in this study was 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 et al. (2001, p 103) note that communication is achieved through all modes separately and, at the same time, together. The writer takes a heuristic view of all communicative modes while analysing situations. Elaine Yakura has considered "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 refers to how texts are disseminated, such as printed book, CD-ROM, or computer application. Mode

refers any organized, regular means of representation and communication, such as, still image, gesture, posture, speech, music, writing, or new configurations of the elements of these (Kress et al. 2001, p 184). 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." 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 touch upon most aspects of multi-modal discourse and its analysis.

METHODOLOGY

The Method of MDA

Multimodal discourse analysis (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:

1. Ideational (the topic of the content)
2. Interpersonal (relationship between the participants)

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' vs '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.

Limitations of MDA:

- It is time consuming
- It is often very technical
- Language aspects tend to be neglected
- Wider social issues may be involved

Communication modes as data types for analysis include:
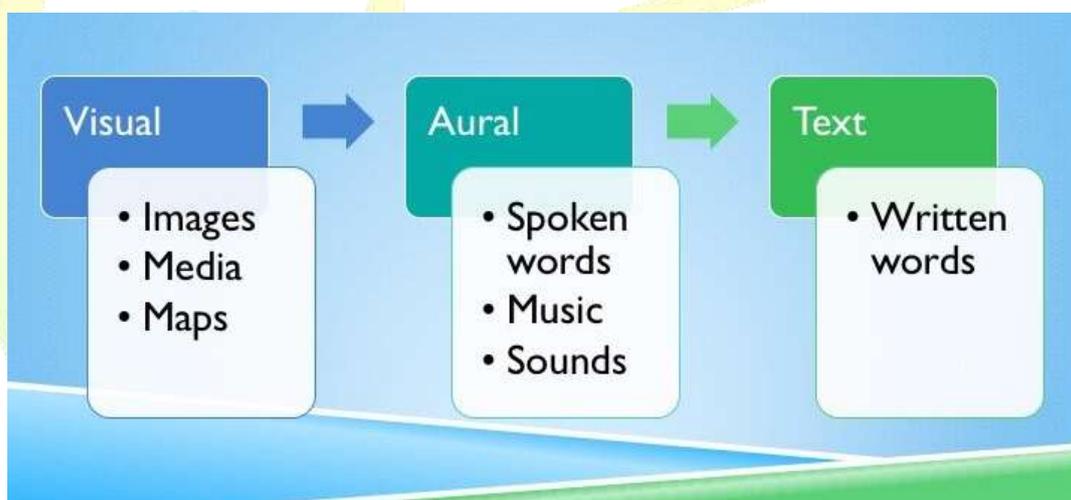
Process:

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

Approaches to MDA

Systemic functional multimodal discourse analysis (SF-MDA) The SFL theory is developed by Michael Halliday. He viewed language as a social system, a meaning making process. SFL has developed a systemic functional grammar (SFG) to represent the meaning making potential of language and to provide a model that helps to understand discourse. These dimensions are used in systemic functional multimodal discourse analysis.

645

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 SFL and CDA.

Multimodal interactional analysis

It early times, it focussed on analysis of conversation 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.

Principles for analysis

This study is based on the principles proposed by Biber and Conrad"s (2009) book, *Register, Genre and Style*. The YouTubers" type of discourse can be described as a register by itself, a specific type of register to be precise." Biber and Conrad define register as "a variety associated with a particular situation of use (including particular communicative purposes)" (2009:6)

According to Biber, the first principle is 'situational'. It is about the physical context, the actual time and place, register, genre etc. the study will follow the following pattern suggested by Biber and

Conrad (2009) - participants, relations among participants, channel (mode), production circumstances, setting, communicative purpose and topic.

DATA ANALYSIS

Content: This video teaches only one word. The word 'plethora' is rather uncommon and difficult for students to understand. The language content consists of 4 examples of the use of the word 'plethora' in sentences. the teacher begins by giving the meaning of the word. 'Plethora' means something in excess or in large quantity. The meaning given is "in excessive quantity". Considering that the students belong to the primary section, the word 'excessive' may also be unknown to them. However, it is followed by a simpler explanation – 'a lot of something'. The sentences are as follows:

- The little boy was so happy when he received a plethora of candy on Halloween night.

- Despite a plethora of videos offered by the video store, Jason always rents the same videos over and over again.

- Because the restaurant staff accidentally created a plethora of food, many diners received free meals.

- Of course, a plethora of spiders isn't good, unless you really like spiders.

a. Participants: The participants are scattered worldwide and can be anyone. The teacher or instructor speaks in the video. He reads out the sentences that contain the new word 'plethora'. He is not visible anywhere in the video. His voice is heard.

b. Interaction: Except the chat section, there is no scope for interaction among participants.

c. Mode: It is a spoken mode. The teacher reads out the sentences that appear in print along with his voice. However, it is not face-to-face communication, but communication through electronic media.

d. Production circumstances: Natural speech is spontaneous and unplanned. There may be repetitions or corrections. In the videos, 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.

e. 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 2009:44). The setting might be a studio or simply any room where the videos is filmed. The setting for each sentence is seen through the images and animation.

In the video, the time and place are not shared. The video is made as per the convenience of the teacher and it can be watched anywhere and at any time by anyone.

f. Topic: The topic is an introduction to the new word 'plethora'.

g. Purpose of communication: The purpose is academic. The teacher tries to teach the new word 'plethora' to the students as part of the vocabulary development process.

Multimodal Features

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

Visual features: The main screen, or the title screen is related to the first sentence. It depicts the boy looking eagerly at the bundle of candies near him. He has a happy expression on the face. he is standing right in front. The position of the boy is significant. It keeps the main

attention of the viewers on the boy who experiences joy.

The video begins directly and abruptly with the display of the word and its meaning on the whiteboard, which appears automatically. There is no physical intervention of the teacher for the purpose of writing. An image of the teacher is there in the left-hand corner of the screen in a circular frame, only in the initial shot. As in several videos, a common practice is to greet the viewers and introduce the topic, followed by explaining the objectives of the lesson. This kind of greeting, in addition to self-references like "Today, I am going to…." help to build a rapport with the learners. However, these features are totally absent in this video.

Every sentence is supported by a still or animated image. The first image consists of a background of Halloween party. A small boy standing in the front jumps and screams (yippee) happily when he gets a 'plethora' of candies. The candies are in a bundle near the boy. He flings up both hands and waves them in joy. As the teacher utters the sentence, it gets types serially on the screen. The second sentence follows immediately. The change in the image is rather quick. More time should have been allowed for young learners to note the details in the image. This could

help fixation of the image and its meaning consequently.

The second sentence shows a video shop with three customers lined up. The image changes after a while, with Jason pointing out to a particular movie. The views guess that he must be Jason because he is the one who points out at something, as if wanting it. There is no animation here, but a change in the images. Two images are used in order to support the sentence. There is nothing special about the setting; the shopkeeper appears to be practical and busy with his work.

For the third sentence, the image seems to be rather weak. Only three characters are seen with a small table in the front. There is very little food on the table and two glasses. Two characters look like diners while the third, seems to be one of the staff, dressed in a uniform. He is uncovering a dish which has a big lid, as if he is about to offer a special surprise. There is nothing in their pose or expression that shows joy, satisfaction or surprise. Only for a fraction of a second, the full screen is visible. Then the typed sentence, which is too long, covers the image, which puts one off. The image is still.

The fourth sentence has no connection with the image. It would be more correct to say that no image is provided to support the sentence. A blackboard is seen on the

right back side of the screen with a list of words on it. The image of the teacher pops in for a minute in the lower right-hand corner and disappears quickly. Then somebody runs fast on the screen – only the feet are visible and the video stops as abruptly as it had begun.

Audio- features: throughout the video, a soft music with a rhythm plays in the background. The boy in the first image shouts the word 'yippee' to express his joy. The reading of the teacher against the background music is the only other audio element.

The music suits the topic and is pleasant. The joy is nicely expressed through the word 'yippee'. All the sentences are clearly read by the teacher and the utterance of each word matches the timing of the word being typed on the screen. The teacher reads with medium speed and with correct pauses, pronunciation and intonation. However, there is no time gap between the sentences. they succeed each other too rapidly, allowing no time to digest the meaning of the sentence, the associated image or the use of the new word.

Pragmatic features

The aim was to make the viewers, that is primary level students, understand the meaning of the word 'plethora'. There is no feedback system to know whether the objective was realized. If the students understand the meaning as well as use the word, it can be said that the video fulfils the locutionary and illocutionary speech acts.

DISCUSSION:

Animated images are suitable for primary students. The background images help fixation and better understanding of the concept. Images must be in accordance with the sentence content. For example, in the first image, the candy is in a bundle and cannot be seen. In the third sentence, instead of showing the setting in the restaurant, the kitchen setting would have been apt, with huge pots filled with a plethora of food. After the images for 3 sentences, viewers expect something in the last sentence but they are disappointed.

Actual use of the new word in various contexts is essential and reveal the usage of the word in sentences. Examples are the best way of introducing vocabulary. One more example did have scope without much affecting the time planned for one word.

The text should not cover the image. It should be positioned in a different way. the other alternative is to keep the image on the screen for half a minute and then start typing the text.

The sentences followed each other very rapidly in succession. The speed must be slowed down to suit the primary level students. Animation makes the videos more attractive. However, there must not be too much of animation that will distract the attention of learners.

Soft music in the background does not distract learners; on the contrary, it is pleasant to the ears. The scream of joy by the little boy at the precise moment adds to the effect positively. More such sounds can be applied.

The examples were apt and correct. There were no spelling or grammatical errors. The duration of the video is five minutes, just right for one single word. The attention span of primary level students is not too long, hence, it is a good idea to include just one word in one video.

CONCLUSION

From the multi-modal discourse analysis of the selected YouTube video for self-learning of vocabulary for primary students it can be concluded that YouTube videos are a good way for teaching learning English vocabulary at the primary level. They must be prepared by experts not only language experts but expert technicians so that more attractive animations and image scan 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. They add personal touch to the videos. The teacher can also suggest some home assignment. A reference to the next video can ensure participation so that students will learn more vocabulary.

650