Unveiling the Art of Food Vlogging: A Multimodal Discourse Analysis of Food Review Vlogs

Nadine Emirey H. Lacsina
Effat University, KSA
nlacsina@effatuniversity.edu.sa
nadinemireylacsina@gmail.com

Abstract

Vlogs are among the videos with the biggest viewership on different social media sites, which led to academic interest and analyses. This study aimed to identify the speech acts used by popular food vloggers in their food vlogs and investigate how the various modalities they used create a perlocutionary effect on the audience in the form of reactions. Furthermore, the study looked into how the modalities observed shape the discourse of food vlogging. This descriptive study explored the genre's apparent unique discourse community through the food vlogs of Mikey Chen, Mark Wiens and Luke Martin in Binondo, Philippines. The corpora which included specific videos and vloggers were chosen through purposive sampling. The linguistic aspect of the data was analyzed by categorizing the utterances into different speech acts, whereas the types of multimodalities present in each of the vlogs were also examined. Results of the study revealed that food vlogs share a common goal of describing food and enticing the curiosity of viewers. It was also found that food vlogs thrive in modes and languages that are exclusive to the food vlogging community. Food vloggers encourage interaction via the comment section. Moreover, food vlogging has a unique genre with specific lexis that includes food, culinary and complimentary terms. Most importantly, food vloggers and viewers in specific food vlogs contribute to the expanding food vlogging community that has its unique discourse which is then shaped by its distinctive modalities and structure.

1. INTRODUCTION

Advancement in technology has created endless opportunities for sharing new information, developing new ideas, and widening networks. In a study conducted by Mueller (2014), he stated that in the advances in technology and networking, people could create and publish their works from different models of media production. Technology has enabled people to use various means to disseminate information, whether written, spoken or both, which come in the form of videos. Linkletter et al. (2009) emphasized, "the influence of online video is so strong". Hence more and more people subscribe to platforms such as social media, which primarily host

videos. Similarly, Nadkarni and Hofmann (2012) stated that social media is a way for people to present themselves to belong in society and to be updated on what is happening around them. Furthermore, these means are carried out through media platforms that people use to gain networks, learn things, and establish relationships.

The number of people subscribed to social media has been increasing daily, making it easier to access news, fashion, food, and events in a more comfortable setting (Osgood, 2017). In this sense, simple videos emerged to become vlogs—a new word formed through combining two genres—video and blog. Vlogging, as defined by Warmbrodt (2007), is similar to text blogging in terms of its instructiveness; however, vlogging gives a more personal, intimate, and empathetic interaction than text blogging, which words and images can only aid.

As the number of social media platforms increases to accommodate more and more videos produced and shared, there is also an increasing number in the kind of genres that these videos are classified into. Blackshaw and Nazarro (2004) gave a general classification of the genres of the different videos uploaded on different platforms: (1) products (2) services and (3) personalities. Furthermore, these categories can still be classified based on the creator's purpose mainly to advertise, teach, evaluate, and review, express and many more. With a huge number of videos uploaded in the video-sharing community, it opens numerous opportunities for researchers to examine the different genres within the video itself and the other genres in terms of different modes such as language, graphics, shots, frames, and structures that are exclusive to each of the videos in every platform.

2. LITERATURE REVIEW
2.1. YouTube is the Leading Video-Sharing Platform

One of the increasing media platforms where people freely share their insights, stances, perspectives, and ways of life is YouTube. YouTube is the leading and preferred social media platform to host videos of different genres and allows video-content creators to interact with their subscribers (Shariffadeen & Manaf, 2017). According to Mullen & Wedwick (2008), "anyone who accesses YouTube can search the large database of videos for his particular need, or a user can set up an account and save selected videos to that account." In the same way, Mueller (2014) stated that video sharing and distribution on YouTube is prevalent for people to communicate their desires to others. The prevalence of video sharing and viewing on YouTube opens opportunities for people to create their own accounts and make their own videos. According to Hilderbrand (2007), the design of YouTube can be compared to television where people can watch one video and turn to another video while clicking a few display buttons on the screen. Hilderbrand (2007) further added that YouTube is a user-friendly website with a simple design and mechanism that facilitates its usage. With this, users can easily watch, stream, and share their own videos (Hu et al., 2019; Zitouni et al., 2021). Moreover, Youtube's feature of allowing its users to connect to Instagram is one advantage that the video platform has (Briliana et al., 2020)

2.2. Food Vlogging on YouTube

As Youtube progressed, more and more people have been drawn to vlogging. Vlogging is also a way to share information, opinions, and preferences about a particular subject that prove to be a powerful tool to influence consumers on their decisions and purchases (Briliana et al., 2020)

Food vlogging is the successor of food blogging. With the advent of technology and introduction of new modalities, food vlogging has taken the spotlight and its presence made its way to Youtube. In fact, food vlogs are even more accessible to everyone because of Youtube.

According to Briliana et al. (2020), food vloggers are opinion leaders, and whatever information these leaders have affect consumers' perception of what they deem true. Aside from this, food vloggers also can influence consumers' choices through electronic word-of-mouth (Zainal et al., 2017). Food vlogs can be likened to food blogs that provide information about a person's dining experience and descriptions of restaurants and their menus. A food vlog is unique because it is in the form of a video on top of other modes that can be used, such as texts, sound effects, images, angles and shots. Since the emergence of food vlogging, greater value and respect have been given to this genre by food consumers (Mumuni et al., 2019). Consumers acknowledge the fact that food vloggers are experienced in the field of food tasting, restaurant hunting and hopping. Food vlogs benefit travelers the most. These travelers look for recommendations through social media because they trust reviews of people who already tried the products and wish to try themselves (Oliveira et al., 2020; Munar & Jacobsen., 2014).

2.3. Changes in Conveyed Meaning

As the leading video-sharing platform, YouTube is home to numerous vloggers who stream videos for personal, business, or interactional purposes. In a study conducted by Warmbrodt (2007) on the vlogging community, he emphasized that technicalities mainly influence the vloggers' preference for the format of their work. As such, videos, as their most preferred format, allows for the combination of audio and images. Videos easily capture their real-time reactions in the form of linguistic and visual features. Linguistic features consist of formulaic expressions, rapport, profanity, and parody which all require the use of words; on the other hand, visual features refer to paralinguistic which include gestures, facial expressions, and other non-verbal elements (Jukic, 2018). These features make Youtube videos appear relatable and reachable. Given the features mentioned earlier, videos are known to be multimodal as they use more than one mode of communication. Multimodality has become more prevalent due to the presence of the internet and social media. As simple as it may be, multimodality refers to the use of two or more modes of communication to be able to effectively convey a message. As what Jones and Hefner (2012) underscores, "In speech aural elements such as the pace, rhythm and the tone of your voice, as well as visual elements such as your gestures, facial expressions and body language all contribute to the message that you send." (2012, p.50). These vlogs provide viewers an array of different modes, such as language, audio, image, frame, gaze, gesture and many more. Researchers have started carrying out studies on how these different modes present in videos contribute to how the meaning of videos is formed. As Jewitt and Price (2012) point out, the current use of video platforms gives very good opportunities to examine the process of meaning-making and establishment of interaction which are both dependent on the different modes that comprise it. With the many modes that can be studied and looked into among the videos, the multimodality approach was used to carry out research of this nature.

2.4. Multimodality in Video Analyses

Most research on videos, images, and posters is carried out through a multimodality approach as it looks into not just language but also other aspects that may impact meaning making. As stated by Kress (2009), multimodality is strongly dependent on social semiotic theories of communication. However, using the multimodality approach not only caters to semiotics but also extends interpretations of language and its meaning as manifested through the different modes of representation. In the same way, Jones and Hefner (2012) supports Kress (2009) by stating that aside from linguistic elements, visual elements in the form of gestures and facial expressions contribute to how a message is understood. Subsequently, Jewitt and Price (2012) emphasize that multimodal analysis does not only consider language as the sole means for making meaning but also considers other modes such as gaze, gesture, and posture.

In a study by Riboni (2017) on makeup tutorial videos, the study used a multimodal approach in analyzing the "makeup guru phenomenon", which showcased the extent of the knowledge of the gurus in terms of language use, rhetorical structure, and social practice of makeup tutorials. Results of the study showed that there are elements only exclusive to makeup tutorials and that these tutorials are formulaic at the textual level. The rhetorical level is fixed, while the social practice level represents makeup as a necessity. On the other hand, a multimodal study conducted by Pratiwy and Wulan (2018) on the Dettol TV advertisement revealed that the audio, gestures, and visual elements make the message of the advertisement more meaningful to the viewers. On a similar note, Castillo Acosta (2018) studied the multimodal aspect of Gambino's music video. Results showed that through the different modes analyzed, such as pose, facial expression, social distance, shot, and objects, mass and social media discourse represent hatred, stereotype and racism that unconsciously shapes the viewers. A study conducted by Yinyuan and Yanfen (2018) on a promotional video of Hangzhou revealed that the modes present in the video contributed to images representative of China. These representations include history, culture, landscapes, life and people. Through these modes, China can be known throughout the world as how the representation of images was done. In another multimodal study by Jukic (2018), the multimodal features of Youtube videos were analyzed in the form of an influencer in the name of Jenna Marbles. Findings of the study revealed that the chosen Youtuber used both linguistic and visual features. Linguistic elements that emerged from the study were formulaic expressions that helped the vlogger communicate better with her subscribers. In contrast, visual elements, such as the vologger's gestures and facial expressions, added to the meaning she wanted to convey.

Simply put, multimodal studies reveal the true meaning of corpora through modes. These modes, when analyzed, reveal the entirety of the meaning that the corpora convey to its viewers, readers, and audience. However, researchers may also opt to consider specific modes that will be subjected to analysis. In this study, the language used and modes such as frame, shot, and generic stage will be magnified.

Studies on food vlogs are very scarce. Only studies about blogs, advertisements and tutorials under a different genre are readily available; however, a study on Filipino vlogs, in general, was conducted by Fuerte (2017), and the results revealed that "pakikipagkapwa" is the dominant ideology of Filipino vlogs. In addition, the researcher emphasized that vloggers can proliferate stereotypes using the modes such as signs and symbols present in the videos. Another recent study by Briliana (2020) on how food vloggers influence millennials revealed that the food vlog reviews vloggers make are comprised of 1) the food and drinks available, 2) recommendations based on the restaurant's menu, 3) affordability, 4) guide on how to consume food and 5) restaurant location and hours of operation. In addition, the study revealed that food consumers value the reviews made by food vloggers, as they see the vloggers as experts in the field. Research made specifically on food vlogs is still not enough to discuss the discourse that the food vlogging community has. This research may be one of the first to focus on the discourse of food vlogging.

2.5. Typology of Speech Acts

The Speech Act Theory was pioneered by Austin (1962) and was reworked by Searle (1975). For this research, Searle's Typology of Speech Acts was utilized. Speech acts are frequently associated with illocutionary acts, part of a three-stage level. The locutionary act, the actual utterance, is the first facet. The illocutionary act, the classification of utterance, is the second facet. Perlocutionary act, the effect on the hearer, is the third facet. Illocutionary acts can be classified into five: representatives, directives, commissive, expressive, and declaratives. Representatives express the beliefs of the speaker. Directives express commands and let the

hearer do something. Commissive expresses future action. Expressive expresses attitudes or states. Declaratives express changes in the state of affairs.

2.6.Television And Genre Analysis

Iedema's (2001) Level of Analysis for Television and Film Genre focuses on the technical aspect of video analysis. There are six major key elements examined: frame, shot, scene, sequence, generic stage, and work as a whole. The frame is described as the key shot representing visual subjects. A shot is defined as an uncut take that only changes if actions such as zooming, panning or tracking occurs. A scene contains more than one shot that exists within the same time and place. A sequence consists of more than one shot that focuses on a certain character or subtopic. Generic stages are generally parts of the video that can be found at the beginning, middle and ending with each genre having particular stages. Lastly, a work as a whole can either be in the genres of 'factual' or 'narrative', which in turn are woven themes of social-cultural, industrial-economic, and symbolic-mythic dichotomies.

2.7.Research Objectives

The objectives of this research were reflective of the problems that the research sought to answer. The objectives of this study were (1) to identify the speech acts used by the vloggers in the food vlogs (2) to discover how texts draw on a range of modalities to create perlocutionary effect, and (3) to know how modalities shape the discourse of food vlogging.

3. METHOD

With the aim of unveiling the kind of discourse food vlogging has, this research utilized descriptive qualitative analysis. To achieve a multimodal analysis of the study, an overarching set of theories were used to explain the discourse aspect of the paper which designates the "how" of the objective—how the discourse of food vlogging is formed. This research drew its bases upon two theories: (1) Speech Act Theory (Searle, 1969) and (2) Iedema's Level of Analysis for Television and Film Genre (2001).

The theory of speech act was used in the study to examine the kind of illocutionary force that the food vloggers use which will help understand its eventual effect or perlocutionary act on the viewers. In addition, the discovery of the illocutionary acts in the vlogs will show the kind of discourse the food vloggers have. The speech act theory covered the language aspect of this research while Iedema's (2001) framework for film and television analysis was used to analyze the following modes: frame, shot, generic stage, and work as a whole. Using the multimodal approach to discourse analysis, consideration of the modes mentioned above is of key importance in understanding the discourse that food vlogging has as these modes also contribute to the entire meaning of the actual video. The food vlogs were examined on the bases of the modes that were mentioned to account for the multimodal approach that this research adhered to.

3.1.Corpora of the Study

The corpora of the study were three food vlogs, which are downloaded from YouTube. The corpora were chosen through Purposive Sampling. The food vlogs are chosen based on the following criteria: (1) place (2) time (3) number of views. These food vlogs all talk about Filipino food to facilitate cultural nuances and at the same time avoid certain biases. In addition, the three food vlogs were uploaded by three different vloggers who were chosen based on the following criteria: (1) number of subscribers (2) length of experience in vlogging (3) cultural orientation. The vloggers were: Mark Wiens, Mikey Chen, and Luke Martin all of whom have been experts in the field of food vlogging.

4. RESULTS
4.1.Types Of Illocutionary Acts

Through a rigorous analysis of the transcriptions of the food vlogs, the findings revealed that there are only four (4) types of illocutionary acts that food vloggers use when doing food review vlogs. These illocutionary acts are: representatives, directives, commissive, and expressive. There were no utterances that could be classified under declarations based on definition, declarations must be uttered by persons in authority. In the case of the vloggers, none of the chosen vloggers is in authority to declare for changes in the state of affairs to happen.

Table 1 shows the specific distribution of entries (utterances) under the four different types of illocutionary acts. The illocutionary act with the highest percentage and received the biggest number of utterances is representatives at 60.97%.

The second illocutionary act with a bigger number of utterances and a percentage of 26.40% is expressive. The number of utterances under expressives is almost half the number of representatives. Mainly, the expressions made by the vloggers under this illocutionary act are complements on the food such as delicious, very good, good, and pretty good. In addition, words of appreciation like thanks and thank you very much were also frequently used by the vloggers. Also, greetings to viewers and restaurant crew such as hello, hi, bye, and bye-bye were used in all the vlogs analyzed.

The third illocutionary act that received a percentage of 11.94% with 86 utterances is directives. Most of the utterances made under this type were questions like "can I have one?", "can I buy one?" and commands like "let's try that!", "look!", "check this out!". All these utterances pertain to the food subject to review and the audience watching the vlog.

The fourth illocutionary act with the least percentage of 0.69% is commissive. The vloggers rarely used commissive in their vlogs. The only commissives they mentioned were "we'll see you again soon", and "we'll see you on our next video" which technically are at the end of the video. All the vloggers considered in the study never used commissive at the beginning and middle parts of the videos. All those 5 utterances were made towards the end of the videos.

TABLE 1. Distribution of Illocutionary Act

Type of Illocutionary Act	Number of Utterances	Percentage
Representatives	439	60.97%
Directives	86	11.94%
Commissives	5	0.69%
Expressives	189	26.40%
Total	720	100%

4.2.Vlogger's Use Of Illocutionary Acts

To further the analysis of the speech acts, all three vloggers' utterances were analyzed and the frequency of illocutionary acts were noted. Table 2 shows the individual distribution of utterances on each illocutionary act by each vlogger. Among the three vloggers, Mikey Chen has the most number of utterances made at 284 while Luke Martin has the least with 186. Regarding the specific illocutionary act, for representatives, Mikey Chen has the greatest number of utterances under this type at 168 while Luke Martin has the least with 114. For directives, Mikey Chen also got the most number of directives at 40 while Luke Martin has the least with 12. For commissive, Mikey Chen has at least 3 while the two other vloggers have 1

each. For expressives, Mikey Chen has the most number at 73 while Mark Wiens has the least with 58.

TABLE 2. Individual Vlogger's Illocutionary Act Distribution

Name of Vlogger	Representatives	Directives	Commissive	Expressive	Total
Mikey Chen	168	40	3	73	284
Luke Martin	114	12	1	59	186
Mark Wiens	157	34	1	58	250
Total	439	86	5	190	720

4.3.Utterances Of Vloggers And Their Perlocutionary Effect On Viewers

All the different speech acts were present in the food review vlogs. Locutionary act, the act of speaking, producing sounds, and contextualizing meaning, were present in the vlogs. Illocutionary acts were observed in the form of representatives, directives, commissive, and expressive. The analysis of the illocutionary acts paved the way for the discovery of the perlocutionary act—the effects of the illocutionary acts on the audience— subscribers and viewers. As stated by Frobenius (2014), "vlogs are predominantly mediated through spoken language and gestures, whereas comments are, by design, mediated through written language (p. 206). To unveil the perlocutionary act, the comment sections of all three vlogs were considered and reviewed further. Comment sections of all three videos from the three vloggers garnered almost a thousand comments respectively. Only the recurring comments from the sections were considered to facilitate the analysis. To get the general categories of comments, recurring comments were analyzed. After getting the recurring comments, two general categories emerged: positive feedback from the subscribers/viewers and positive feedback from Filipinos. The videos chosen were food review vlogs on Filipino food in Binondo. Aside from the two general categories, there are subcategories for each general category that emerged from the analysis.

The first general category is positive feedback on the videos, which is composed of six (6) subcategories: words of appreciation, venue suggestions for the next food vlogs, attention on vlogger gestures during food tasting, viewers' urge to try the food and visit the place of the shoot, and viewers' expression for continuous subscription to the channel. These subcategories emerged from combining all the recurring comments across all the comment sections. These subcategories came from comments from viewers and subscribers of different nationalities, genders, ages, and cultural backgrounds. Comments referred to as words of appreciation all talk about how thankful the viewers are to the vloggers for creating such a food review video. In addition to thanking them, the viewers also appreciate the vloggers for uploading videos of their food vlogs from time to time. Viewers not only thank the vloggers but also suggest places that the vloggers may consider visiting and doing a vlog about. Most places that viewers mention are their home countries or the countries they are staying in. In the case of vloggers' enticing gestures, viewers commented on how vloggers react and move when tasting food. In fact, viewers can easily evaluate the food presented based on the reaction and gestures of the three vloggers. Among the vloggers, Mark Wiens' comment section was filled with comments on his "lean" gesture when eating. Many comments on the sections were about the viewers' urge to try the food shown in the logs. This may have been the most frequent and common perlocutionary effect on the audience. With all the representatives, directives, and expressives that the vloggers used, the audience has become very eager to try the food shown in the vlogs. Subsequently, after expressing their urge to eat, most viewers also express their support for the three vloggers. Most comments indicated that the viewers have been following the vloggers and have been watching all the video uploads of the vloggers.

The second general category is positive feedback from Filipinos, which is composed of three (3) subcategories: suggestions on other Filipino foods to try, corrections on some of the food descriptions, and warm welcome from the Filipino community. Filipinos dominated the list among all the viewers who commented on the respective vlogs. The large number of Filipino comments posted was due to the fact that the videos were all about Filipino food. Filipinos expressed their excitement and happiness through the comments by suggesting other signature Filipino dishes for the vloggers to try and other places in the Philippines to visit. Moreover, since Filipinos are very familiar with the food shown, they were keen on the food descriptions so there were some posts correcting an incorrect food description; however, these corrections were done well and were accompanied by neutral or friendly emoticons.

Furthermore, the warm welcome of Filipinos was seen in all the comments sections. Most Filipino comments were "welcome to the Philippines" and "thank you for visiting the Philippines". Consequently, Filipinos also showed their love for the vloggers by saying, "we love you" and "Filipinos love you".

With all the categories of comments presented, the vlogger's utterances are able to create perlocutionary effects on the viewers. To further look into the specific perlocutionary effects, refer to the table below.

TABLE 3. Perlocutionary Effects of Vlogger Utterances on Audience

Positive Feedback on Videos	Positive Feedback from Filipinos
Words of appreciation	Suggestions on other Filipino food to try
Venue suggestions for the next food vlogs	Corrections on some of the food descriptions
Attention on vlogger gestures during the food tasting	A warm welcome from the Filipino community
Viewers' urge to try the food and visit the place of shoot Viewers' expression for continuous subscription to the channel	
Words of appreciation	
Venue suggestions for the next food vlogs Attention on vlogger gestures during food tasting	

4.4. The Genre Structure Of Food Review Vlogs

A thorough analysis of the corpora revealed that food vlogs are structured into three general parts: (1) the introduction and expectations, (2) the food tasting and food reviewing, and (3) the summary of highlights. Simply put, there are five moves in food review vlogs: introduction-expectations-food tasting-food reviewing-summary of highlights.

The first part, which is usually found at the beginning of the video, follows a generic sequence. First, the food vlog title is projected along with the place of visit. Second, the vlogger covers the introduction for important content elements in the video. Among important elements that are introduced include the vlog or channel of the vlogger, the vlogger, and the place featured in the video. At times, a quick history of the place featured is also provided. Third, the vlogger gives a summary of the itinerary, food to try, and places to visit to set viewer expectations.

The second part, which is the larger bulk of the video, is the actual food tasting and food review. The general sequence of this section includes a walking trip to traditionally famous areas that serve equally known food. Vloggers either use the help of local food vloggers or contacts to

show them to these areas and guide them through the food review, or gather information from travel forums and sites to determine recommended places to go and food to consume. The actual food review is done through food tasting, food description and corresponding food price features.

The last part, which is the shortest, follows a usual sequence that includes a summary of food, place (at times weather) and experiences, an invitation to visit the featured places and try the food, and lastly, a nudge to promote the vlogger video and channel. In the summary section of the last part, descriptions of food, place and experiences are kept short and focus only on the highlights. Shortly after highlights and personal preferences are emphasized, recommendations follow. Lastly, before bidding viewers goodbye, the common push to promote vlogger videos and channels involves encouraging viewers to maximize basic YouTube functions such as liking and commenting on videos and subscribing to the channel.

4.5. Language and other Modes of Food Vlogs

Language and modes shape the discourse of the food vlogging community. On the language aspect, through the speech acts, findings revealed that vloggers frequently use representatives which are purely descriptions of the food that is subject to taste test, and directives which are commands and questions that vloggers raise so viewers can answer them through the comment sections, commissives which are vloggers ways of encouraging viewers to look forward and anticipate the next food vlogs to be released, and expressive which are mostly about the vloggers' judgment on the taste of the food.

In the creation of meaning, several other points of analysis were considered aside from language. These points are modes that help add meaning to the corpora. In this study, videos were chosen as corpora. Videos are not restricted to words and language alone; other modes compose videos and make them a unique genre. This study utilized the framework of Iadema (2010), which enumerated modes such as frame, shot, scene, and sequence that showcase the uniqueness of television and film genres. Furthermore, for the purpose of this study, a few additional modes were also considered: gestures, gaze, sound, and image.

For further elaboration, the modes are discussed individually on the basis of how it was shown in all three videos. The modes include general aspects such as (1) gesture, (2) gaze, (3) sound, and (4) image. Image, on the other hand, is specified into four other sub-areas such as (4.1) frame, (4.2) shot, (4.3) scene, and (4.4) sequence. Some finer details show the contrast between vloggers and their respective videos. However, each category and its subcategories are generally expressed similarly by vloggers.

The first general aspect is gestures. Vloggers have unique gestures that viewers commonly associate with them and are perceived as manifestations of their characters. Gestures such as balling hands into fists and pursing lips when food is delicious or beckoning to the camera as a sign of invitation and a close companionship relationship with viewers were observed. Common gestures during food tasting that denote good-tasting food include nods of approval or closing/widening eyes.

The second general aspect is gaze. The study results show that vloggers maintain eye contact with the viewers of the video. While some vloggers maintain eye contact at all times to establish a more conversational type of communication with their viewers, it is also common for some vloggers to keep their gaze moving from the food to the camera in a more presentation-type review.

Another general aspect that contributes greatly to the meaning and intent delivered by vlogs is sound. While some vlogs use any form of music that reduces dead air in scenes where not

talking is involved, some vlogging channels generally use music that creates an immersive experience by choosing music that is culturally appropriate to the places featured in their videos.

On the aspect of the image, frames, shots, scenes and sequences are utilized to shift focus, give emphasis, and/or show variety. Generally, frames are designed to introduce vlogs, vloggers and their channels. The elements found in frames are logos which are either shown first without video content, or are shown in a manner that blends with shots of the place featured in the vlog. Moreover, close-ups of food and food preparation are shown before the vlogger starts tasting and reviewing the food. Some vloggers utilize time-lapsed shots of places, use slow-motion effects in food preparation shots, and upward or side panning shots to show environments where featured food is found. In cases of multiple dishes of food served, shots are framed to show the vlogger seated on a table with the food placed between the camera and the vlogger.

Subsequently, scenes and sequences that are generally identified as a string of shots put together in many different creative ways depending on the vlogger's purpose are edited in a manner that sustains interest. Some vloggers use panning or zoom-in styles of transitioning from one shot to another to create seamlessness in montage scenes; vloggers generally opt for simpler methods of stringing shots of food, tasting and vlogger descriptions and food review as the focus is shifted from editing to content. Sequences, on the other hand, are generally held together by traveling usually done on foot or through transports unique to the place featured in the vlog in an effort to frame the non-cuisine cultural aspect of the said places.

5. DISCUSSION

Starting with the language aspect, representatives were uttered the most on the discussion of illocutionary acts—scoring a percentile of 60.97%. This finding is supported by the fact that most of the utterances made by the vloggers were statements about the food and the place where one can get it and the actual descriptions of the food (60.97%) that they have reviewed and tried. The descriptive nature of such utterances resulted in the high occurrence of representatives of illocutionary acts.

Among the findings that made similarities among the three vloggers apparent include the frequent use of words of appreciation and greetings. This result is supported by the study of Frobenius (2014) which stated that most of the available vlogs on YouTube always start their videos with common greetings such as hello and hi.

In addition to the language aspect of the analysis, commissive was not encountered in all three vlogs at the beginning of the video. This finding strongly suggests that food vloggers do not use commissives except when referring to future videos or sequels of videos they are showing at the time of speaking. All in all, the individual distribution of the illocutionary act still points out the earlier findings that most utterances fall under representatives while commissives had the least.

On the viewer end of the study, the results of analyzing categorized comments showed that vlogger utterances do have perlocutionary effects on the viewers. This finding only shows that food review vlogs can affect peoples' views on food, interests in food, and peoples' taste buds.

Considering other elements such as gestures, the vloggers are associated with manifestations of their characters such as bodily movement such as Mark Wiens' tendency to lean when he appreciates his food. This result is supported by a study on vlogging by Frobenius (2014), which emphasizes that body gestures are influenced by how the vloggers talk facing the camera. In addition, he stated that gestures in monologues like vlogging revealed that the presence of an imagined audience or listener affects the frequency and types of gestures done

by the vloggers. Through these simple gestures, viewers get more points of relation from actual descriptions made by vloggers.

Another element considered is gaze. According to the result revealed in the study of Frobenius (2014), vlogger's usually gaze at the audience which helps establish rapport. In the study, vloggers were observed to maintain eye contact with the viewers of the video.

Conclusively, while descriptives are common and dominating, directives push viewers to comment or ask, commissive encourages subscription and anticipation of sequel vlogs, and expressive detail vlogger judgment, these given illocutionary acts show the kind of utterances the food vloggers produce to create a discourse exclusive to food review vlogging alone. Simply put, on the language aspect, these food review vlogs are dominated by food descriptions and judgments of food taste. Combined with additional elements analyzed in the study including gestures, gaze, sound and image, a package of these modes creates a discourse that is exclusive to the food vlogging community.

Meanwhile, on the technical aspect, a pattern emerged post analyses of the three vlogs. There is a standard food vlog title presentation and a view of the place visited. Then, an introductory part follows, introducing not only the video, the vlogger and the channel, but also the featured place. This finding is parallel with the study conducted by Forbenius (2014), which stated that vlogs often start with credit, music, and image. These entities are regularly shown on a channel that is similar to TV series with a title shown on screen. All the vlogs analyzed contained all these entities, which also served as the identity of their unique respective channels. In addition, toward the end of the video, a summary of the experience is shown. This apparently encourages viewers to visit the place visited, try food that were featured, and engage viewers by inviting them to leave a 'like' on the video and subscribing to the vlogger's channel. This finding is similar to the results found in the study of Frobenius (2014), which revealed that a big number of vlogs contain closings, which are sequenced as follows: closing components, terminating components, arrangements, and acknowledgments. In the case of the corpora, all these aspects were visible and observed.

As for the finer details of the technical rudiments of the vlogs, images, frames, shots, scenes and sequences were used to control viewer attention. Editing of the aforementioned created moods for the vlogs, shifted attention where the focus was intended, and projected variety. Vlogs that utilized culturally appropriate music framed the desired form of projected atmosphere for their respective vloggers. The vlogs presented their logos or initial frames either as 'logo only' or 'logo along with the featured place'. This allows viewers to set initial expectations. Subsequently, shots are utilized to put certain elements at the forefront of viewer attention. Often, close-ups of food and food preparation are shown to hook viewers before the actual vlogger tasting and review. Video speed was also utilized differently; some vloggers time-lapsed their shots of the place they visited to show the busy atmosphere, while others slowed their videos about food preparation to give emphasis to food preparation steps. Lastly, there were instances where vloggers framed shots where the multiple dishes of food were placed between the camera and the vlogger to emphasize variety.

6. CONCLUSION

This current study aimed to identify the speech acts that food vloggers used, examine the perlocutionary effect of the utterances of the vloggers on the viewers, and reveal the discourse of food vlogging by looking to the genre and structure of the food vlogs.

Findings revealed that food vlogging has its own language, through speech acts, and has its own genre, through the examination of different modes. Put together, these aspects create a discourse community that is only exclusive to food vlogging alone.

On the language level, with the aid of the speech act theory, the food review vlogs observed all three aspects of speech acts: locutionary, illocutionary, and perlocutionary acts. Locutionary act in the form of vloggers' act of speaking. This act makes food vlogging far more appealing and trustworthy to food consumers, as they can see exactly the vloggers trying the food themselves. As a result, more food consumers become more dependent on vloggers for their expert reviews. Meanwhile, illocutionary acts were in the form of representatives, directives, commissive and expressive. As expected for a food review vlog, representatives dominated the analysis of the vlogs because giving statements is the lifeline of food vloggers. No dead air was noticed all throughout the vlogs and this was possible because the vloggers spoke all the time. For the perlocutionary act, this was observed in the form of comments made by the viewers upon watching the vlogs. Through the comments made by the viewers, it can be concluded that food vloggers have their own ways of drawing responses from their viewers, may it be in the form of asking questions and suggestions. In addition, it is concluded that despite the fact that both vloggers and viewers do not interact face to face, viewers are able to develop an increased attachment to the food vloggers of their choice by mainly watching their food vlogs. This attachment was evident in the comment section and led to a well-grounded conclusion that food vlogging shapes the way food consumers or viewers see food reviews. Viewers put so much respect on the food vloggers and would consider going to the place to give the food a try or at least give the food a try when they are within the area. This was a good manifestation of a perlocutionary act.

On the genre level, the vlogs clearly showed a structure that is unique to this genre. The vlogs all have a three-part structure with a total of five moves. These structures are introduction and expectations, food tasting and reviewing, and summary and highlights. These patterns are not observed in other vlogs pertaining to different genres such as a tutorial, meet and greet, presentations and the like. The structure that emerged is exclusive to the food vlogging community. Hence, prospective food vloggers or novice food vloggers may consider using this structure in order to attract more viewers to watch their vlogs. Although they may add more structures to their vlogs, considering the three parts that emerged in this study would prove useful to start a food vlogging career, as the vloggers examined in this research are considered to be the pioneers of food vlogging.

On the level of the different modes such as gesture, gaze, sound, image, shot, frame, and sequence, all the results of the analysis of these modes point to the general findings that the food review vlog as a genre comes with modes that are uniquely exclusive to it. Not only are the structures different but also the modes. In food vlogging, gesture, gaze, and shot were the most unique ones in this genre of vlogging. Through the close examination of modes, vloggers capitalize on their facial expressions when trying food because these seem to draw attention from the viewers. Aside from this, how the vlog was shot or recorded was also a factor to consider. Hence, vloggers do not just focus on the actual presentation and tasting of food, but also make it a point to take good shots and videos of the place where they try the food to make the vlog a complete package to gain more subscribers.

Through the findings of this research, the discourse community of food vlogging can be summarised as follows: (1) it has a common goal of describing and trying the food, enticing/attracting people to try the food, and encouraging people to visit places (2) it has its own modes and language that are exclusive to the food vlogging community (3) food vloggers encourage viewers to communicate with them through the comment section (4) it has a unique genre (5) it has specific lexis—food, culinary, and complimentary terms (6) it has individuals who contribute to the discursive expertise—these are the actual vloggers, viewers and other equally knowledgeable persons whom the food vloggers meet in specific places to help them describe, look for, and try the food.

The studies on vlogs have not received much attention in the research community yet. Hence, studies related to this research were scarce. However, the findings of this research open doors for a lot of opportunities for doing studies on digital platforms by examining their genre, language, and multimodality.

6.1. Recommendations

This study may serve as an opportunity to encourage discourse analysts to look into food vlogging as a field of study due to the increasing attention this field of vlogging gets from different social media platforms. However, it must be noted that this study only looked at some of the famous food vlogging pioneers, hence more research must be done to be able to examine closely the multimodality and features of metadiscourse that are present in food vlogging to thoroughly analyze how unique the discourse of this genre of vlogging is.

REFERENCES

Austin, J. L. (1962). *How to Do Things with Words*. Oxford: Clarendon Press.

Blackshaw, P., & Nazzaro, M. (2004). Consumer-Generated Media (CGM) 101: World- of- mouth in the age of the Web-fortified consumer. Retrieved September 1, 2022 from http://www.brandchannel.com/images/papers/222_cgm.pdf

Briliana, V., Ruswidiono, W., & Deitiana, T. (2020). Do Millennials believe in food vlogger reviews? A study of food vlogs as a source of information. *J. Mgt. Mkt. Review,* 5(3) 170 – 178. https://doi.org/10.35609/jmmr.2020.5.3(5)

Castillo Acosta, M. F. (2018). A Multimodal Discourse Analysis of Childish Gambino's Music Video "This is America". *Open Journal for Studies in Linguistics*, *1*(2), 57-72

Frobenius, M. (2014). Audience design in monologues: How vloggers involve their viewers. *Journal of Pragmatics.* 72:59-72

Fuerte, D. M. (2017). Vlogging sino ka man: On vlogs and created filipino identities. [Unpublished Undergraduate thesis]. University of the Philippines, Manila, Philippines.

Hilderbrand, L. (2007). YouTube: Where Cultural Memory and Copyright Converge. *Film Quarterly*. 61:48-58.

Hu, M.; Chen, M.; Li, M.; Yin, Z. (2019). Meet the media characters from another culture: Influence of ethnocentrism on parasocial interaction. *Psychology of Popular Media Culture*, v. 8, n. 2, 170-179. https://doi.org/10.1037/ppm0000168

Iedema, R. (2001), Analyzing film and television: A social semiotic account of Hospital: An Unhealthy Business. *The Handbook of Visual Analysis*, 183-204.

Jewitt, C. & Price, S. (2012). Multimodal approaches to video analysis of digital learning environments. *In The 26ᵗʰ BCS Conference on Human Computer Interaction* 26 (pp.1-4)

Jones, R. and Hafner, C. (2012) *Understanding Digital Literacies: A practical introduction*. New York: Routledge

Jukic, A. (2018). An analysis of multimodal features of Youtube videos. University of Zagreb. Repozitorij Filozofskog fakulteta u Zagrebu' at University of Zagreb. https://core.ac.uk/download/299375324.pdf

Kress, G. (2009). *Multimodality: A Social Semiotic Approach to Communication*. London: Routledge Falmer.

Linkletter, G. & Dooley J. (2009). The Choking Game and YouTube: A Dangerous Combination. *Clinical Pediatrics*. 49:274.

Mueller, B. (2014). Participatory culture on YouTube: a case study of the multichannel network Machinima. Retrieved July 6, 2022 from http://www.lse.ac.uk/media@lse/research/mediaWorkingPapers/MScDissertationSeries/2013/msc/104-Mueller.pdf

Mullen, R., & Wedwick, L. (2008). Avoiding the digital abyss: Getting started in the classroom with youtube, digital stories, and blogs. *Clearing House: A Journal of Educational Strategies, Issues and Ideas,* 82(2), 66-9.

Mumuni, A. G., Lancendorfer, K. M., O'Reilly, K. A., & MacMillan, A. (2019). Antecedents of consumers' reliance on online product reviews. *Journal of Research in Interactive Marketing,* 13(1), 26-46. https://doi.org/10.1108/JRIM-11-2017-0096

Munar, A. M., & Jacobsen, J. K. S. (2014). Motivations for sharing tourism experiences through social media. *Tourism management*, 43, 46-54. https://doi.org/10.1016/j.tourman.2014.01.012

Nadkarni, A., & Hofmann, S. G. (2012). Why do people use Facebook? *Personality and Individual Differences*, 52(3), 243-249.

Oliveira, T., Araujo, B., & Tam, C. (2020). Why do people share their travel experiences on social media?. *Tourism Management*, 78, 104041. https://doi.org/10.1016/j.tourman.2019.104041

Osgood, N. A. (2017). Finding the purpose of tasty videos according to social media audiences. *Elon Journal of Undergraduate Research in Communications*, 8(2), 88-95

Pratiwy, D. & Wulan, S. (2018). Multimodal discourse analysis in dettol tv advertisement. Retrieved September 1, 2022 from https://knepublishing.com/index.php/Kne-Social/article/view/1932/4320

Riboni, G. (2017). The Youtube makeup tutorial video: A preliminary linguistic analysis of the language of "makeup gurus". *Lingue e Linguaggi*, 21, 189-205. Doi:10.1285/i22390359v21p189

Searle, J. R. (1969). *Speech Acts.* London: Cambridge University Press.

Shariffadeen, T. S., Manaf,. A. (2017). Following Islamic Reality Show Personalities on Twitter: A Uses and Gratifi cation Approach to Understanding Parasocial Interaction and Social Media Use. *Intellectual Discourse,* 25, (Special Is), 637-659. https://journals.iium.e du.my/intdiscourse/index.php/islam/article/view/1064

Warmbrodt, J. (2007). An exploratory study on videoblogger's community. Scholar's Mine. Retrieved August 22, 2022 from https://scholarsmine.mst.edu/cgi/viewcontent.cgi?article=5574&context=masters_theses

Yinyan, Y. & Yanfen, Z. (2018). A multimodal discourse analysis of the promotional video of Hangzhou. *English Language Teaching*. 11(10), 121-131

Zainal, A., Harun, A., & Lily, J. (2017). Examining the mediating effect of attitude towards electronic words-of mouth (eWOM) on the relation between the trust in eWOM source and intention to follow eWOM among Malaysian travelers. *Asia Pacific Management Review*, 22(1), 35-44. https://doi.org/10.1016/j.apmrv.2016.10.004

Zitouni, M., Al-Traif, H., Zemni, B., Mohammed, O. S., & Aljasser, M. (2021). Utilization of YouTube to improve the pronunciation skill of Saudi learners in translation departments. *Asian ESP Journal*, *17*(2), 133-154.

About the Author

Dr. Nadine Lacsina *is a Senior Lecturer at Effat University in Jeddah, Saudi Arabia. She completed her Master of Arts in English as a Second Language in 2017 and Doctor of Philosophy in Language Studies in 2022 from De La Salle University. She has been teaching English to ESL/EFL students for over twelve years now. She has delivered several lectures on instructional materials development and pedagogical practices. Her research interests include Discourse Analysis, Semiotics, Second Language Acquisition, Sociolinguistics, and Applied linguistics.*