



Investigating formulaic expressions in an informal YouTube broadcast interview video

Talitha Anindya Tsani^{✉1}, Galuh Kirana Dwi Areni¹

¹Universitas Negeri Semarang, Indonesia

Article Info

Article History:

Received on 5 July 2025

Approved on 20 July 2025

Published on 17 August 2025

Keywords: Discourse analysis; Formulaic expressions; YouTube; Podcast; Spoken Discourse

Abstract

This study investigates the use of formulaic expressions in spontaneous informal communication found in a YouTube video titled 50 Cents: 50 Minutes You Can't Afford to Ignore in a podcast interview format. The focus of this study is to analyze the types, frequencies, and pragmatic functions and meanings of the use of formulaic expressions in conversational contexts based on Celce-Murcia's Communicative Competence theory (2007) and discourse analysis approach. The identified expressions are then categorized based on four types such as routines, collocations, idioms, and lexical frames. The findings show that collocations are the most frequent expressions in the video followed by lexical frames, routines, and idioms. These results indicate that formulaic expressions, especially collocations and lexical frames, play an important role in maintaining the flow of interaction so that it runs smoothly, naturally, and effectively in the context of informal interaction. This study also highlights how the implications of using formulaic expressions to convey ideas or messages in an easy-to-understand, effective, clear, and interesting way can strengthen the emotional bond between the speaker and the interlocutor or audience. The findings are also in line with several previous studies that have examined the dominance of the use of formulaic expressions in different contexts. This study contributes to a deeper understanding of the use of verbal language through digital platforms, provides insight into the use of formulaic expressions in real-life spontaneous interactions, and can be used as a reference for future linguistic research with a focus on digital communication or discourse.

[✉] Correspondence Address:
B3 Building FBS Unnes
Sekaran, Gunungpati, Semarang, 50229
E-mail: talithat125@students.unnes.ac.id

INTRODUCTION

The development of technology in this modern era has brought about major changes in various aspects of human life. Technology is now no longer considered as a mere complement to life, but has become an inseparable part of human daily activities that can help facilitate all activities carried out (Manongga, 2022). Not only help in terms of work efficiency and communication, technology also plays an important role in the world of education and science. One of the impacts of the advancement of information technology is the emergence of various digital platforms and social media which are now not only used as a means of entertainment, but also as a source of information, data, and even objects of scientific study. One of the most popular platforms today is YouTube. This video-based platform allows users from all over the world to access, share, and create content freely and quickly. According to Maharani and Budiarti (2022), YouTube is now used as a public media to convey various types of information, ranging from entertainment, education, to social and cultural discourse. Because of its effectiveness and efficiency, YouTube is not only a communication tool but also one of the main sources in the development of language and culture considering its global reach so that it can help everyone access any kind of content from all over the world.

The uniqueness of YouTube lies in the form of its content which is multimodal, which can combine audio, visual, text, expression, and visual elements that can display body language (Benson, 2016). This makes YouTube a very rich source of data for linguistic research, especially in the field of oral discourse analysis or verbal communication. Videos such as vlogs, interviews, podcasts, TED Talks, or casual discussions available on YouTube display the use of language in a very natural and unscripted context (Cooper, 2023). This is what makes YouTube different from written data sources or formal recorded scripts, because the language used in it truly reflects how people speak in real life.

One interesting linguistic phenomenon to analyze in real communication on YouTube is the use of formulaic expressions. This study focuses on formulaic expressions through the lens of Celce-Murcia's (2007) Communicative Competence theory. To understand language use in the context of spoken discourse, communicative competence is a crucial concept. The theory of communicative competence that was developed Celce-Murcia in 2007, offering a comprehensive approach encompassing six main components, one of which is formulaic competence, which serves as the primary reference in the analysis of formulaic expressions. Formulaic competence refers to fixed language structures that are readily available without the need for grammatical reprocessing and frequently appear in everyday interactions by native speakers. In this study, formulaic competence is categorized into four categories: routines, collocations, idioms, and lexical frames. These four categories assist researchers in the process of identifying, classifying, and explaining how these expressions function in building effective communication.

By using YouTube as a data source, researchers have access to a variety of formulaic expressions used by native speakers in very diverse situations. One relevant example to analyze is an interview video titled *50 Cent: 50 Minutes of Advice You Can't Afford to Ignore* uploaded by Bryan J. Roberts on September 25, 2021, on his YouTube channel. This video contains an informal conversation between the interviewer, Bryan, and the guest star, 50 Cent, who is a famous American rapper and actor. The interview was conducted in a relaxed yet meaningful style, with topics discussing 50 Cent's life journey, career, and personal views. The conversation in this video reflects the use of language in a natural context that is full of formulaic expressions, both in the form of collocations, idioms, routines, and lexical frames.

The reason for choosing this video as the object of research is because of its authentic, relevancy, and takes place in real communication that occurs spontaneously without using a script to control the interaction. This makes the video very suitable for linguistic analysis, especially to identify, classify, calculate the most used expression, and reveal the contextual meaning of the used expressions in real communication in the theme of informal podcast.

Many studies have investigated the use of formulaic expressions in spoken English, especially in real-life communication contexts such as classroom interactions. For example, Khoiriyah & Mujiyanto (2022) and Sa'adah et.al, (2024) found that collocations and lexical frames were the most frequently used types of formulaic expressions during oral exchanges in informal language learning environment and formal classroom context. These expressions were characterized as familiar, easy to understand, and frequently occurring in everyday conversations. This study is also in line with a study conducted by Neno & Agustien (2016) regarding the use of formulaic expressions in students' interactions in Timor State University. The research found that collocations and lexical bundles were the most frequently used expressions in interactions. This dominance also indicates that students tend

to be more familiar with expressions that are easily understood directly compared to those with idiomatic meanings. Because of its highly contextual use, the types of formulaic expressions can also vary depending on cultural background, register, and social situation. Linguistic studies that focus on formulaic expressions also provide an in-depth understanding of how language is used in everyday life practices. This is relevant in the digital era, where communication no longer only occurs face-to-face, but also through online media such as YouTube, Zoom, Instagram, and other platforms.

In summary, while previous research has acknowledged the distribution of formulaic expression used in every day interactions, especially in classroom context, few studies have specifically examined how these expressions function in informal, unscripted conversations such as those found in YouTube-based interviews or discussions. The current study aims to address this gap by analyzing the types and frequencies of formulaic expressions in authentic YouTube interviews, with a focus on how they are used naturally in spontaneous communication using discourse analysis approach through the identification of formulaic expressions based on Celce-Murcia's Communicative Competence theory (2007) to classify the formulaic expressions based on their types and pragmatic functions.

Through this analysis, this study not only aims to describe linguistic phenomena descriptively, but also seeks to contribute to the understanding of the use of formulaic expressions in real communication so that later it can be practiced directly by language learners as a medium to communicate to each other, especially English, so that the language delivered can sound more fluent and more natural like a native speaker and help them to effectively convey messages or ideas by combining the use of the types of formulaic expressions that will be revealed through this study. This study does not focus on the pedagogical aspect or foreign language learning, but purely as a linguistic study of the spontaneous use of language in digital media.

METHODS

The approach used in this study is a descriptive qualitative approach aimed at exploring the use of formulaic expressions that naturally occur on digital platforms. The object of this study is an interview video titled *50 Cent: 50 Minutes You Can't Afford to Ignore* from Bryan J. Roberts' YouTube channel, which features Bryan as the interviewer and the famous American actor and rapper, 50 Cent, as a guest star. The video discusses his career path, life experiences, and other messages, rich in motivational advice, and is therefore intended for audiences seeking life motivation. This approach was chosen because of its relevance in uncovering the meaning and function of language in real interactions through narrative descriptive presentation rather than numerical data (Creswell & Creswell, 2017). This research is accompanied by a discourse analysis method, which aims to understand the function of language in constructing identity, influencing audiences, and creating meaning in specific contexts (Gee, 2014). The use of this method is highly relevant given the selection of videos rich in linguistic elements that can be used as data and processed using discourse analysis.

This research focuses on the use of formulaic expressions, encompassing four categories based on Celce-Murcia's (2007) theory of Communicative Competence: routines, collocations, idioms, and lexical frames. These expressions are part of spoken language spoken by native speakers, making them suitable references for examining videos containing spontaneous and natural communication and aiding in the data identification and classification process.

Data collection was conducted online from March to July 2025 via YouTube. The videos were 50 minutes long and produced transcripts totaling 8,089 words, excluding time stamps. The research location was flexible, utilizing digital space that authentically represents the phenomena under study (Moleong, 2017). Data collection techniques included documentation and non-participant observation, conducted by watching the videos in their entirety and transcribing the utterances, either manually or with the help of an automated transcript available on YouTube to streamline the transcript extraction process.

The data was then analyzed using a combination of content analysis and discourse analysis. Content analysis was used for identification and classification, while discourse analysis was used to interpret the contextual function and meaning of utterances appearing in video interactions. The analysis was conducted through stages such as transcript extraction, identification of utterances that included formulaic expressions based on definitions and examples from communicative competence theory, classification, calculation of frequency of occurrence, and interpretation of meaning. The results of the analysis were compiled in the form of a classification table and a summary of findings presented descriptively, accompanied by diagrams to facilitate the visualization of the communication

functions of the identified expressions. Examples include the routine expression "..., right?" to request confirmation, the collocation expression "at that point" to describe a specific time, the idiomatic expression "hang out" referring to casual activities with close friends or family, and the lexical frame expression "I think..." used to express personal opinions.

The analysis process also followed the stages of data reduction, presentation, and conclusion drawing as described by Miles et al. (2018). Data reduction was carried out by selecting the most relevant and meaningful data. The presentation was arranged in a descriptive narrative format accompanied by a frequency diagram of the use of formulaic expressions. Conclusions were drawn by linking the findings to the theoretical framework used.

Data validity was ensured through investigator triangulation techniques. According to Flick (2018), investigator triangulation involves comparing multiple perspectives from researchers and other observers to re-examine data validity to ensure scientific accountability and minimize bias. In this study, the researchers involved colleagues and competent internal parties. One of the participants who contributed to the validation of this study was an Indonesian citizen who had resided for five years in a country where English is native. The informant was deemed to have a sufficient understanding of the use of formulaic expressions in everyday communication. Through discussions, a more specific classification of formulaic expressions was obtained, which was then classified according to their respective categories.

By implementing the above methods, this research is expected to contribute to linguistic studies, especially regarding the use of formulaic expressions through digital.

FINDINGS AND DISCUSSION

This section is specifically created to present the results of the analysis of the use of formulaic expressions in a video podcast entitled *50 Cent: 50 Minutes You Can't Afford to Ignore* uploaded on Bryan J. Roberts' YouTube channel. Based on the identification results, 1359 expressions were found (including expressions that were repeated) consisting of four categories of formulaic expressions based on the Communicative Competence theory by Celce-Murcia (2007), namely routines, collocations, idioms, and lexical frames. These four categories appear with varying frequencies and are based on the personal experience of 50 Cent, as a guest star who is present to provide advice and motivation through his life experiences. In general, the researcher found that the speaker in this video podcast uses formulaic expressions as a means to convey ideas, maintain interaction with the interlocutor, and bind the audience emotionally. To deepen the discussion, the findings and percentage results of each category will be described in the following chart presented in Figure 1 and subsections:

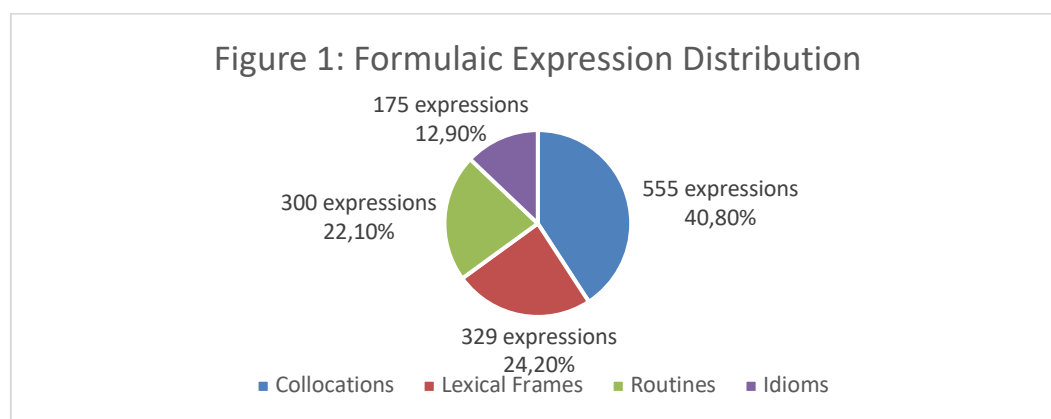


Figure 1. Formulaic Expression Distribution

Collocations (555 expressions / 40.8%)

In the analysis conducted from a total of 1359 expressions, 555 of them were considered as collocations that appeared most frequently in the video. Collocations are pairs of two or more words that appear naturally and generally appear together when communicating and usually contain meaning in a certain context (Brezina et al., 2015). The form of collocations can consist of adjective

+ noun, verb + noun, verb + adjective, and many other forms. Collocations play an important role in maintaining native-like fluency and effective communication (Bui, 2021; Durrant & Schmitt, 2009).

The results of the study found that there were two examples of collocation expressions that often appeared in the video. The two examples include "at that point" which was said 9 times, and "kind of" which was said 7 times during the video. These two examples are collocations that are commonly found in English conversations and have an important function to maintain smooth communication and provide certain meaning in the message to be conveyed.

The expression "at that point" is a collocation in the form of a prepositional phrase that can be used to indicate a specific time in communication. In the video, an example of the use of this expression can be found in 50 Cent's speech at minutes 46.17-46.25 as follows:

"But it's not a part of the success. I'm being successful, doing things that are a lot different from what I was doing at that point."

In this example, the use of the expression "at that point" refers to a specific time in 50 Cent's past life that describes him still trying to do different things before reaching the peak of his success now and that is not what successfully led him to success. The function of using this expression is to describe 50 Cent's personal development and make it easier for the audience to capture the story or personal reflection of 50 Cent's life. This explains that what 50 Cent did in the past is not part of the success he has achieved now so that anyone who has the same experience as 50 Cent will be motivated to achieve success in the future in the same way. Another example of collocation expression is use of the expression "kind of" which can also be found in the following quotes:

"...But, that kind of faded after a while, once I got into the groove of doing things..."

The use of "kind of" at minutes 5.24-5.30 serves as a marker of doubt about the feelings experienced by 50 Cent when telling his experience. He describes his experience of experiencing a sudden change in complex feelings and is hesitant to assume that the feelings have changed for sure. The use of this expression invites the audience and Bryan as the host to interpret these feelings together and bind the audience emotionally to his story.

Both expressions are used by 50 Cent not only to convey life experiences, but also to invite his audience and interlocutors to feel and understand the emotional process and self-reflection that he is experiencing so that they can take positive messages and make 50 Cent's life experience a motivation to understand their own feelings. This also strengthens that collocation expressions play an important role in forming meaningful communication and building emotional connections with the audience.

Lexical Frames (329 expressions / 24.2%)

Researchers managed to identify 329 expressions that are included in the lexical frames expression and make it the second highest frequency after the use of collocations. Lexical frames are defined as lexical patterns that are repeated and can accommodate many variables in the language structure (Ang & Tan, 2019). Lexical frames are different from lexical bundles which are continuous sequences of words, lexical frames allow for additions in a phrase but still maintain its original structure (Gray & Biber, 2013). Celce-Murcia (2007) includes examples such as would you like to..., it is important to..., I think..., can I have..., and many other examples of phrases in her research that allow for the addition of information or have a flexible structure to complete a sentence with the intention that the message to be conveyed can be received well, maintain the smoothness of discourse, and make it easier for the interlocutor to follow the flow when interacting.

Two expressions were found that often appeared during the video, including the expression "I think..." which was said 13 times and "and then..." which was said 9 times. The expression "I think..." is commonly used when the speaker wants to convey a personal opinion in a subtle way that does not seem pushy and allows room for interpretation for the audience. The use of this expression in full context can be found in 50 Cent's speech at minutes 6.17-6.27 as follows:

"And when you have someone's training that really don't, I think the neighborhood had them,"

In this example, 50 Cent uses the expression "I think..." to give his personal opinion about the surrounding environment and its influence on someone who is in the training process. The use of this

expression can help convey opinions that may contain sensitive discussions so that it can soften the way it is conveyed so that it can be received by the ears of the audience or the person he is talking to well. In addition, 50 Cent also opens up a dialogue space or invites the person he is talking to respond to his opinion by providing arguments that may be contrary. Meanwhile, the use of the lexical frame expression “and then” can also be found in 50 Cent’s speech as follows:

“...and then anyone who’s extremely focused could be considered ruthless at points because of they get in the way where I’m doing, I just knocked them out the way.”

Those examples uttered by 50 Cent at minutes 8.17-8.28 shows the function of the expression “and then” as a connector for the narrative being conveyed. This expression is used to sequence an event or situation in his life story. Based on the example above, the use of “and then” is used to be a connector between the previous idea and what happens afterward and keeps the narrative connected naturally. In addition, the expression also functions as a temporary pause that can make it easier for the audience to digest the previous idea before moving on to the discussion afterward.

In conclusion, the use of lexical frames in the video shows that the effectiveness of interaction depends on how the speaker conveys the message through flexible and familiar language patterns so that the audience or interlocutor does not have to struggle to digest the message being conveyed. With various functions as transitions between ideas, as discussion controls, and also as tools to build connections with the audience, lexical frames are a very important part of formulaic expressions that can support the effectiveness of communication, especially in the context of discourse.

Routines (300 expressions / 22.1%)

A total of 300 expressions (22.1%) have been successfully identified by researchers and are categorized as Routines. Routines can be interpreted as expressions that are used repeatedly during communication and can be spoken directly without the need to construct new sentences grammatically so that it can reduce a person's cognitive load (Pawley & Syder, 1983). According to Nattinger & DeCarrico (1992), routines include various kinds of expressions in the form of direct expressions such as greetings, apologies, or gratitudes. This expression is often used in informal conversations as a way to maintain smooth communication, strengthen the connection between speakers, and ensure that the other person understands the intent conveyed and also help develop discourse competence, making it easier for someone to communicate in a real context. (Celce-Murcia & Olhstain, 2000).

One of the most frequent forms of routines in the video is the expression “..., right?” which was found 18 times, uttered by both 50 Cent as a guest star and Bryan J. Roberts as an interviewer. This expression functions as a confirmation marker, namely when the speaker wants to make sure whether the listener agrees with the previous statement. In addition, the use of “..., right?” also creates a more relaxed, intimate atmosphere and encourages deeper two-way interaction. An example of this usage can be found in the following quote by 50 Cent at 0:46–0:59:

“So, it just makes me not look at what would be the bigger situations in front of me or harder circumstances don't seem so rough to me no more because I've already been up against the toughest things, the biggest obstacles. If you, I think, take this for granted, right?”

In this sentence, 50 Cent wants to emotionally invite Bryan and the audience to agree with his opinion that many people often take things for granted without understanding the struggle hidden behind them. The use of “..., right?” in this context serves to strengthen the sense of togetherness and connection with the audience.

In addition, two other forms of routines that are also often found in this video are the expressions “You know what I’m saying?” 15 times, and “You know what I mean?” 11 times, all of which are said by 50 Cent. Although the forms are different, both of these expressions have a similar communication function, namely to ensure that the listener understands the essence of the message being conveyed. This phrase usually occurs after 50 Cent has explained a particularly complex or emotional life experience or personal view, and is used to grab attention, get the listener involved, and establish the impression that he is speaking directly and honestly. For example, at 46:29–46:40, 50 Cent says:

"I had a kid that I was mentoring for a little while that, the kid was a bad kid, bro. He was just a bad kid, all right? But he was in love with everything about all the wrong things about me. You know what I'm saying?"

In those examples of quote, he is sharing his personal experience of being a mentor to a child who is actually emulating the negative side of himself in the past. The phrase "You know what I'm saying?" is used to ensure that Bryan and the audience can understand the message he is conveying, while also inviting them to see the situation from his perspective more deeply. Overall, the use of routines in this interview not only serves as a strategy to keep the conversation lively and smooth, but also as a tool to build a strong connection between the speaker and the listener. Routines such as "...right?", "You know what I'm saying?", and "You know what I mean?" reflect 50 Cent's way of delivering messages with a distinctive charm that can strengthen communication between speakers and indirectly with his audience.

Idioms (175 expressions / 12.9%)

The expression that appeared the least frequently during the video interview between Bryan J. Roberts and 50 Cent was the idiomatic expression, with 175 expressions that were successfully identified. Idioms are utterances in the form of words or phrases that have hidden meanings. According to Biber & Conrad (1999) and Fernando (1996), idioms are defined as expressions that have idiomatic meaning and whose original meaning cannot be predicted from their lexical structure. However, unlike the previous examples that have repeated expressions, the expressions that appear in this video are almost dominated by expressions that are unique and only appear once. The researcher found several examples of expressions that are commonly used in the context of everyday conversations that appear in the video such as "hang out", and "get into". To find out the meaning and function of the use of these expressions, the researcher will describe examples of the use of expressions in full context as follows:

36.39-36.51 : "...So, it just was so much success that in both two different directions that they didn't see it until later when things slowed down a little bit and they got a chance to hang out with each other."

20.19-20.29 : "Now, when you get into street life or into that fast lifestyle, they'll tell you that the outcome is going to be you dead or in jail."

The quote with the expression "hang out" has two meanings, one of which means to be suspended in a place. However, seen from the full context, this expression refers to another meaning that is more commonly used, which refers to casual or non-formal activities that are usually done with friends or family. This expression describes communication that tends to be more relaxed and casual.

Meanwhile, the use of "get into" can be interpreted as involving oneself in a particular activity or situation. In the context of the quote above, 50 Cent uses this expression to explain someone's involvement in street life that may be related to illegal or criminal activities that can cause someone to end up in an unpleasant situation such as ended up in prison.

The two examples of idioms above show how 50 Cent can bring his audience and interlocutor into his life story and experiences which are wrapped in such a way in his distinctive narrative. The use of idioms can also add color to the story and enrich the meaning so that the message conveyed sounds so interesting to anyone who hears it.

Overall, the findings reveal that collocation is the most frequently occurring expression in the analyzed 50 Cent: 50 Minutes You Can't Afford to Ignore video, followed by lexical frames, routines, and idioms. This pattern shows how speakers tend to use casual and informal expressions when conveying their life experiences, making it easier for the audience to grasp the ideas conveyed. These results are in line with research conducted by Sa'adah et al. (2024) and Khoiriyah & Mujiyanto (2022) which found similar findings in the form of the use of collocations followed by lexical frames that dominate the classroom interactions studied. The high percentage of collocation and lexical frames use indicates that speakers choose expressions that often appear in everyday conversation, easy to understand, and able to help convey messages or ideas effectively without making listeners think too long to receive the message.

Therefore, the use of collocation is highly recommended for use by beginner English learners or used in casual contexts in other informal interactions. In addition to showing the dominant use of collocations, these findings also highlight the low percentage of idiom use which is also in line with

research by Neno & Agustien (2016). The study revealed that the percentage of idioms is ranked 4th out of 5 categories of formulaic expressions with a percentage of only 1.6% of all formulaic expressions found. This idiomatic expression is rarely used because when compared to other expressions such as collocations or lexical frames which are considered more common or often appear in everyday interactions, idioms are expressions that have a certain meaning that cannot be directly translated literally so that it can make it difficult for the audience or listener to understand the contents of the message or story being conveyed. Therefore, the use of idioms is widely avoided because it is not an effective choice, especially in podcast discourse whose purpose is to provide advice or motivation that can be directly conveyed to the audience easily such as in the video 50 Cent: 50 Minutes You Can't Afford to Ignore. Although its use is complicated and tends to be avoided, it does not mean that the use of idioms is not important in the study of the use of formulaic expressions. Precisely because of its low of use, it allows many researchers in the future to make it an object of research with a different approach and can find results that may be contrary to this study. Thus, the results of this study not only strengthen the findings of previous studies, but also reveal flexible and effective language practice in communication with diverse contexts.

CONCLUSION

This study aims to investigate the use of formulaic expressions in informal podcasts that occur spontaneously in videos on YouTube using a discourse analysis approach. By analyzing a video titled 50 Cent: 50 Minutes You Can't Afford to Ignore, this study successfully revealed 1359 formulaic expressions that were categorized into four categories based on the Communicative Competence theory by Celce-Murcia (2007), consisting of routines, collocations, idioms, and lexical frames.

The findings of this study found that the use of collocations was the most frequent expression in the video followed by lexical frames, routines, and idioms. This pattern shows that in informal verbal communication, speakers tend to choose expressions that are familiar and easy to understand and commonly appear in everyday conversations so that they are effective for use in conveying messages or ideas. In this study, the use of dominant collocations and lexical frames was proven to maintain the smoothness of discourse, facilitate the explanation and distribution of ideas or messages, and bind the audience emotionally so that they feel carried away by the topic of conversation. On the other hand, the use of idioms is the least frequently used or used by speakers in videos due to various factors such as requiring a long process to be digested by the audience because it has a figurative meaning and is less commonly found in everyday interactions so that it is not effective to use in the context of conveying messages containing advice or motivation.

This study also supports previous studies that highlight the dominance of collocation in real-life spoken interaction. Furthermore, this study contributes to a deeper understanding of how formulaic expressions are used in digital discourse, such as informal interview podcasts. It also provides evidence that the use of YouTube provides linguistic richness and is rich in authentic data.

In conclusion, the use of formulaic expressions plays an important role in forming smooth, natural, and effective interactions or communications, especially in informal verbal interactions. These expressions not only help facilitate the delivery of ideas or messages but also help express emotions, goals, and interactional strategies with the audience and also the interlocutor. Therefore, the researcher hopes that this study can be a reference for future research on the use of formulaic expressions in digital media or platforms and can be used as a reference for English learners and English educators who want to deepen their insight into the use of certain languages or expressions and how to implement them in various discourses and everyday contexts.

FUNDING STATEMENT

This research received no specific grant for any funding agency in the public, commercial, or non-profit-sectors.

REFERENCES

- Ang, L. H., & Tan, K. H. (2019). From Lexical bundles to lexical frames: Uncovering the extent of phraseological variation in academic writing. *3L the Southeast Asian Journal of English Language Studies*, 25(2), 99–112. <https://doi.org/10.17576/3L-2019-2502-08>
- Benson, P. (2016). *The discourse of YouTube: Multimodal text in a global context*. Routledge.
- Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In *BRILL eBooks* (pp. 181–190). https://doi.org/10.1163/9789004653689_017

- Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International journal of corpus linguistics*, 20(2), 139-173.
- Bui, T. L. (2021). The role of collocations in the English teaching and learning. *International Journal of TESOL & Education*, 1(2), 99-109. <https://i-ite.org/index.php/journal/article/view/26>
- Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. *Intercultural language use and language learning*, 41-57.
- Celce-Murcia, M., & Olshtain, E. (2000b). *Discourse and Context in Language Teaching: A guide for language teachers*. <https://ci.nii.ac.jp/ncid/BA52360239>
- Cooper, C. R. (2023). The identification of YouTube videos that feature the linguistic features of English informal speech. *Applied Corpus Linguistics*, 3(3), 100068. <https://doi.org/10.1016/j.acorp.2023.100068>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (4th ed.). SAGE Publications.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL - International Review of Applied Linguistics in Language Teaching*, 47(2). <https://doi.org/10.1515/iral.2009.007>
- Fernando, C. (1996). *Idioms and Idiomaticity*. Oxford University Press.
- Flick, U. (2018). *An Introduction to Qualitative Research* (6th ed.). SAGE Publications
- Gee, J. P. (2014). *How to do Discourse Analysis: A Toolkit*. Routledge
- Gray, B., & Biber, D. (2013). Lexical frames in academic prose and conversation. *International journal of corpus linguistics*, 18(1), 109-136.
- Maharani, D., & Budiarti, E. (2022). Pengaruh media digital & mutu perangkat terhadap kemampuan bahasa pada aud melalui konten youtube. *Jurnal Jendela Pendidikan*, 2(03), 429-434. <https://doi.org/10.57008/jip.v2i03.240>
- Manongga, A (2022). Pentingnya teknologi informasi dalam mendukung proses belajar mengajar di sekolah dasar. *Prosiding Seminar Nasional Pendidikan Dasar*, ejurnal.pps.ung.ac.id, <https://ejurnal.pps.ung.ac.id/index.php/PSNPD/article/view/1041>
- Miles, M. B., Huberman, A. M., & Saldana, J. (2018). *Qualitative data analysis: A Methods Sourcebook*. SAGE Publications.
- Moleong, L. J. (2017). *Metodologi Penelitian Kualitatif* (Edisi Revisi). Bandung: Remaja Rosdakarya.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press.
- Neno, H., & I.R Agustien, H. (2016). The use of formulaic expressions in efl students' interactions. *English Education Journal*, 6(1). Retrieved from <https://journal.unnes.ac.id/sju/eej/article/view/12778>
- Pawley, A., & Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of pragmatics*, 7(5), 551-579.
- Roberts, B.J. (2021). *50 Cent: 50 Minutes You Can't Afford to Ignore*. YouTube. <https://www.youtube.com/watch?v=oO-vjUgJPug>
- Sa'adah, H., Setyaningrum, E. Y., Riski, Y. A., & Afifi, N. (2024). Investigating formulaic expression to improve students speaking skill. *Jurnal education and development*, 12(1), 137-139. <https://doi.org/10.37081/ed.v12i1.5271>
- Schiffrin, D. (1987). *Discourse markers* (No. 5). Cambridge University Press.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press, 110 Midland Ave., Port Chester, NY 10573-4930 (45 British pounds).