Available online at: http://journal.uny.ac.id/index.php/ljtp

LingTera, 10 (2), 2023, 134–152

A corpus-based analysis of discourse marker '*like*' functions in podcasts: A gender study

Zahrotul Nur Azizah Surya Putri[1], **Suray Agung Nugroho**[1]
[1]Universitas Gadjah Mada, Indonesia
*****Corresponding Author. Email: zahrotul.nur.azizah.surya.putri@mail.ugm.ac.id

Abstract
Focusing on the use of discourse markers (DM), this study investigated the use of DM "*like"* in dyadic communication in American Spotify podcasts. This research aimed to understand the linguistic behavior of podcasters in using DM "*like"* during their podcasts and to compare the spoken corpus of female and male speakers collected from the podcast. Corpus analysis assisted by AntConc 3.5.9. was applied to determine (1) which gender primarily used DM "*like"* within the corpus, (2) the significance between the usage of DM *like* and the speaker's gender, and (3) the functions of using DM "*like*" by each gender based on Müller's (2005) taxonomy of DM *like* functions. The analysis portrayed that male speakers prefer to use DM "*like"* more than female podcast speakers, which is in contrast to previous studies claiming that female speakers used more DM "*like"* than male speakers. The difference in results could be attributed to the 17-year gap in the study, which has potentially led to changes in the inclination of using DM *"like"* between genders. However, despite male speakers' higher usage of DM "*like*", there is still no correlation between the use of DM "*like"* and the speaker's gender, aligning with previous research. Lastly, this study revealed that both genders used the same DM "*like*" functions, although the occurrences varied.

Keywords
podcasts; corpus; dyadic communication; discourse markers *like* functions

Citation (APA Style): Putri, Z. N. A. S., & Nugroho, S. A. (2023). A corpus-based analysis of discourse marker '*like*' functions in podcasts: A gender study. *LingTera*, *10*(2), 134–152. https://doi.org/10.21831/lt.v10i1.61245

INTRODUCTION

Language is a powerful tool that not only facilitates communication but also reflects the intricacies of human society, culture and gender dynamics. Society believes that there are certain ways of using language and its features preferred by one gender and stereotypically associated with being feminine or masculine (Wardhaugh, 2015). Lakoff (1975), in his study on language characteristics among genders, assumed that males and females use language in different ways to reflect differences in their ways of thinking, culture, and socialization which become biases among society. One of the different language features shown by males and females includes the choice of discourse markers (Pasaribu, 2017).

Previous researchers believes that discourse markers, which can sometimes be used as fillers, are most likely to be used by female speakers rather than male speakers (Müller, 2005; Gabrys, 2017; Pratiwi, 2020; Pasaribu, 2017). Discourse markers are said to be related to female language features since one of their functions is to show hesitation and make the sentence more expressive and exaggerated, which aligns with female language characteristics (Müller, 2005; Lakoff, 1975). However, discourse markers not only function to show hesitation and exaggeration but also served to make utterances more fluent, clear, and coherent, aiming to improve mutual understanding between speakers and listeners (Gabrys, 2017; Martin-Laguna & Alcon-Soler, 2018; Schiffrin, 1987; Li, 2014). Discourse markers are concrete manifestations of discourse strategy that serve practical

purposes and significantly contribute to smooth and effective oral communication rather than solely reflecting certain gender language characteristics. They assist listeners in grasping the speaker's ideas and intended messages during interactions. Thus, exploring the functions of discourse markers "*like*" between genders in more detail, beyond their function as hesitation and to show expression, is worth analyzing.

Schiffrin (1987) added that discourse markers were extensively applied in essential social interaction (Fung & Carter, 2007; Cahyanti, 2021). The use of discourse markers serves not only as a strategy to enhance communication performance but has also evolved into a communication style, particularly prevalent among the youth community, especially in casual or informal communication (Müller, 2005). Various discourse markers commonly employed in casual communication include *"like," "well," "you know," "I mean"*, etc. Gabrys (2017) claimed that discourse markers were predominantly employed by middle-aged speakers (19-40 years old). Given the extensive use of discourse markers, their types and functions vary concerning the dynamism of language. Therefore, discourse markers have garnered academic attention, with studies focusing on their changes and variations since the late '80s (Drager, 2011; Gabrys, 2017; Pratiwi, 2020; Aijmer, 2022).

Müller (2005), in his study, assumed that the word "*like*" is crucial in the native speaker's communication either as discourse markers or its other function. He also states that the diverse use of *"like"* is a fundamental element of youth talk, which is a complex vernacular that continues to be one of life's fascinating little mysteries. Therefore, the word *"like"* is assumed to have multiple functions, especially when it serves as discourse markers (Gabrys, 2017). Furthermore, over time, the word *"like"* has become a trend in colloquial youth language, particularly in American youth language (Müller, 2005). Taguchi (2002) asserted that young speakers use this word within their utterances without necessarily changing the sentence's meaning. However, each word's usage serves a different function in the utterance, aiding the listener in understanding the speaker's pragmatic meaning and reflecting certain attitudes towards the content being spoken.

Some previous researchers such as Drager (2011), Tolliver (2012), Müller (2005), and Gabrys (2017) specifically focused their study on discourse markers "*like*". Yet, apart from focusing on the word "like" as discourse markers, the aforementioned researchers also explored the other functions of the word "*like"* such as lexical verb, adverb, quotation, etc. In her study on the lemma *"like,"* Drager (2011) focused her study on the phonetic characteristics of the word's usage by New Zealand girls in their utterances. She found that the lemma *"like"* was used in grammatical functions (lexical verbs and adverbs) and discursive functions (quotative, discourse particle, discourse marker, and approximate adverb). However, since the primary focus of the study was on the phonetic variation of the lemma *"like,"* Drager (2011) did not provide an in-depth analysis of the specific function of DM *"like."* Further, Tolliver (2012) also investigated the usage of the word *"like"* and explored its shift in African American English (AAE). In the study, Tolliver revealed that the word "*like"* function in AAE indeed undergoes a shift. Tolliver's (2012) claim aligned with Drager's (2011), stating that *"like"* was used as a verb, quotative, approximate adverb, discourse marker, and discourse particle focuser.

Furthermore, Müller (2005) delved into the usage of the word "*like"* used by native and non-native English. His study assumed that the word "*like"* is crucial for native speaker's communication. Müller (2005) outlined that the word "*like*" can be classified into three groups i.e., (1) discourse markers, (2) quotative, (3) grammatical (verb, noun, adjective, adverb, preposition, conjunction, and fixed expression. Moreover, in his research, Müller (2005) specifically proposed the taxonomy of DM *"like"* functions including *searching for the appropriate expression, marking an approximate number or quantity, introducing example, introducing explanation,* and *marking lexical focus*. Based on Müller's (2005) classification of the word "*like"* function, discourse particle, discourse marker, approximate adverb, and discourse particle focuser mentioned by Drager (2011) and Tolliver (2012) is under the taxonomy of DM *"like"*.

Additionally, Gabrys (2017) conducted a detailed analysis of the use of DM "*like*" in different varieties of English, as found in the International Corpus of English (ICE) and Santa Barbara Corpus of Spoken American English (SBC). She used the above-mentioned Müller's (2005) taxonomy of the discourse marker "*like"* functions and proposed new taxonomies improving

Müller's (2005) framework, namely *discourse restart* and *unclassified instances*. Using the above improved taxonomy, Gabrys (2017) revealed that Irish English speakers primarily used DM "*like"* compared to Canadian, American, British, and New Zealand. Furthermore, she found that female speakers used the DM "*like"* as hedges and intensifiers more frequently than males.

Due to the different categorization of the word *like* among the researchers, this research aims to investigate the functions of the word *"like"* that is served as discourse markers using the improved Müller's (2005) taxonomy proposed by Gabrys (2017) i.e., (1) *searching for the appropriate expression,* (2) *marking an approximate number or quantity*, (3) *introducing example*, (4) *introducing explanation,* (5) *marking lexical focus,* (6) *discourse restart*, and (7) *unclassified instances*. Additionally, since discourse particles and discourse markers are often referred to by different terms, such as pragmatic markers, pragmatic particles, etc. (Drager, 2011; Müller, 2005; Gabrys, 2017), this present study will consistently use the term "discourse markers" to maintain consistency in terminology.

In contrast to previous studies (Drager, 2011; Tolliver, 2012; Müller, 2005; & Gabrys, 2017) which focused on the word "*like*" in general sense, this research will solely focus on the usage of DM "*like*" in American English, providing a comprehensive analysis of its functions used by American speakers. Moreover, since Gabrys (2017) study used the ICE compiled from spoken corpora in the 1990s and the SBC compiled from spoken corpora between 2000-2005, this research seeks to investigate if there have been any changes in the usage of DM "*like"* in American English during 2022-2023, particularly in podcasts. To achieve this, a new corpus is compiled by collecting the most recent American Spotify podcast video transcripts. By doing so, this study aims to present more up-to-date data on the usage of DM "*like"* by American speakers outside the ICE and SBC.

Furthermore, previous researchers have examined the usage of DM in academic interactions (Pratiwi, 2020; Fung & Carter, 2007), essay writing (Ratih & Abidah, 2022), daily interactions, YouTube videos (Cahyanti, 2021), globally available corpora (Gabrys, 2017), or in specific environments such as all-girl high schools (Drager, 2011). In contrast, this research specifically focuses on the usage of DM "*like"* in face-to-face conversations as seen in the podcasts. According to Gabrys (2017), comparing the use of like in podcasts with spoken discourse may reveal additional aspects of the communication situation that are also important in using DM.

Thus, exploring the natural and spontaneous use of DM during communication is an exciting area of study because it allows researchers to examine the functions of DM usage in everyday communication (Clifton et al., 2020). Therefore, to uncover the natural usage of DM "*like*", the researchers intend to analyze the functions of DM "*like"* as used by podcasters (hereinafter referred to as speakers). By exploring podcast conversations, this research aims to elucidate the current usage of DM "*like*" by podcast speakers. Spotify podcasts are chosen as the data source for the spoken corpus because podcasts involve unscripted face-to-face conversations between interlocutors. According to Dohen et al. (2010), communication can be considered face-to-face if the speakers can see each other speaking, observe facial expressions, and interpret body gestures. Based on this definition, podcasts can be considered as face-to-face conversations (Dohen et al., 2010; Clifton et al., 2020).

Moreover, Nurmukhamedov and Sharakimov (2021) assumed that, since 2019, podcasts have become a popular form of media in society and are categorized as natural occurrences of language. Therefore, exploring the usage of DM in podcasts is expected to reveal the usage of DM in real life. To facilitate easy access and data retrieval, the researchers utilized the recorded Spotify podcast videos uploaded by the podcast speakers on their official YouTube accounts.

Furthermore, this study aims to explore the function of DM "*like"* within the corpus and focuses on which gender is more likely to use DM "*like"* during the podcasts. Therefore, once all occurrences of DM "*like"* in the corpus are known, this research intends to determine whether gender has any correlation to the usage of DM "*like*," as reflected by female and male speakers in the Spotify podcast videos. This investigation seeks to determine whether the data will reveal different results compared to previous studies. Lastly, this research is designed to provide a detailed analysis of the functions of DM "*like"* used by podcast speakers in their utterances.

RESEARCH METHOD

This research employed a mixed method, combining qualitative and quantitative approaches to analyze the function of DM "*like*" using Corpus Linguistics data. The corpus used in this research was derived from spoken transcripts of selected American Spotify podcast videos. The corpus was compiled by collecting recorded podcast transcripts from two female-female podcast videos and two male-male podcast videos uploaded on their respective official YouTube channels. The channels were selected randomly using random sampling, considering several predefined categories related to the speakers' social identity that may influence language characteristics (Gaither et al., 2015). These predefined categories included white American speakers, speakers of the same gender, and middle-aged speakers (19-40 years old) (Tipton, 2014; Gabrys, 2017). Two episodes were selected from each of the male-male and female-female podcast accounts, with each episode having a duration of 45 minutes. To facilitate the analysis, the corpus was divided into two sets: female and male, consisting of utterances from female and male speakers in the podcast. Thus, both of female and male corpus consisted of four episodes with a total duration of 180 minutes. Based on the transcripts, the female corpus comprised 36,245 words, while the male corpus comprised 46,150 words.

Firstly, the corpus data is analyzed using the Ant Conc 3.5.9.0 concordancer tool, as the research instrument used in this research. This tool assists in identifying the occurrences of DM "*like"* in the corpus. Additionally, quantitative analysis was employed to determine the correlation between speaker's gender and the tendency of DM "*like"* usage in the utterances of female and male speakers. To assess the correlation of the two variables, this research was utilized the Chi-Square Test, a statistical test developed by Karl Pearson (1857-1936), to analyze the correlation between non-numeric variables, and Cramér's V test developed by Harald Cramér's (1946) to measure the strength of the association between the speakers' gender and the usage of the DM "like" or the effect size (Turhaan, 2020; Kearney, 2017). Through this statistical hypothesis test, the research aims to establish the correlation or significance between gender and the usage of DM "*like"* in the spoken corpus. However, as the data used in this research are derived from two different corpora with varying sizes (female corpus: 36,245 words, male corpus: 46,150 words), it is necessary to normalize the data to enable a meaningful comparison of the frequency of DM "*like"* usage (Biber et al., 1998). According to Rundblad (2015), normalization is essential to ensure an accurate and equitable comparison between two discourses, whereby the two corpora are made equivalent in size (Rundblad, 2015; Biber et al., 1998). Therefore, to normalize the data, the researchers adjusted the data to a base of 10,000 words, and the results are presented in Table 1 in the subsequent section. Following the normalization process, the data becomes comparable and can be described and compared (Rundblad, 2015; Biber et al., 1998).

Secondly, the qualitative method in this study was utilized to analyze the usage of DM "*like*" by female and male speakers. To narrow down the scope of the study, the primary focus is on DM "*like*", especially the function, based on Müller's (2005) taxonomy of DM "*like"* functions by Gabrys (2017). Consequently, instances where the word "*like*" is not served as a DM will not be considered.

FINDINGS AND DISCUSSION

Based on the podcast corpora, the occurrences of the general word "*like"* and those that serve as DM within female and male corpus can be seen in the following table.

Table 1. Normalized Occurrences of the Word *like* in the Corpora

Category	Raw Occurrences		Normalized Occurrence (per 10.000 words)	
	The word *like*	DM "like"	The word *like*	DM "like"
Female Corpus	1507	1108	415.8	305.7
Male Corpus	1979	1451	428.8	314.4

As seen in Table 1, it can be inferred that the occurrences of the word "*like*" in the female and male corpora are slightly different. The results also show that the usage of "*like*" as a discourse marker (DM) is no longer dominated by female speakers. According to the results, out of the 36,245

words collected in the female corpus, the word "*like*" serves as a DM occurred 1,108 times out of 1,507 tokens, equivalent to 305.7 out of 415.8 occurrences in normalized data. In other words, female speakers use 73.5% of the total occurrences of DM "*like*" within 180 minutes of conversation. Meanwhile, according to the data, the male corpus shows that there are 1,979 occurrences of the word "*like*" out of 46,150 words, with 1,451 instances serving as a DM. These occurrences amount to 314.4 DM occurrences out of 428.8 occurrences of the word "*like*" in normalized data. Therefore, it can also be said that within 180 minutes of conversation, 73.3% of the total occurrences of "*like*" are used as a DM. From this perspective, it can be concluded that male speakers produce the word "*like*" more frequently than female speakers do in their natural conversations presented in the podcast.

Further, to specifically discuss the usage of "*like*" as a DM, the data in Table 1 shows that male speakers used 314.4 occurrences of the word DM "*like*," whereas female speakers only produced 305.7 occurrences of DM "*like*" out of the total occurrences. Based on this data, this research concludes that male speakers produced the word "*like*" more frequently than female speakers and used DM "*like*" within their utterances more often during the 180-minute podcast duration. This finding contradicts previous research claims that females produced more DMs than males (Müller, 2005; Tolliver, 2012; Gabrys, 2017; Pratiwi, 2020). In fact, Gabrys (2017) claimed that among American speakers, female speakers used DM "*like*" more frequently, with 67% of the occurrences produced by females, while males only produced 33% (Gabrys, 2017).

In contrast, this study found that in 2023, male speakers in Spotify podcasts used DM "*like*" more frequently than female speakers. Although the difference in DM "*like*" occurrences between the female and male corpora in this study is only 8.7 occurrences per 10,000 words, the data suggests a slight change in the usage of DMs between female and male speakers. Based on this, it can be assumed that the difference in phenomena between the previous study conducted by Gabrys (2017) and this present study is likely due to the 17-year gap between the two corpora. Gabrys (2017) used data from the SBC (Spoken American English) corpus, which captures linguistic phenomena between 2000 and 2005, while this study collected data from the latest linguistic phenomena in 2022-2023. Therefore, it can be assumed that the time gap between the two corpora may lead to different phenomena in using DM "*like*." As Aijmer (2022) stated, language change can occur over time, influenced by social factors, discourse practices, and evolving socio-cultural norms.

Furthermore, the SBC corpus presented in Gabrys (2017) includes conversations in various settings, such as oral interactions, telephone conversations, classroom lectures, and town hall meetings, which cannot be exclusively categorized as casual or formal. Meanwhile, Müller (2005) claimed that DM "*like*" is most likely to occur in casual conversations. Therefore, a different set of conversations within the corpus cannot fully represent the usage of DM "*like*" in casual conversations among American speakers.

However, despite the finding that male speakers used more DMs "*like*" compared to males in 2000-2005, there is still no significant correlation between the usage of DM "*like*" and the gender of the speakers. The chi-squared results can be seen in the following Table 2.

Table 2. The Result of Chi-squared Test

Gender	Raw occurrences of DM "*like*"	Word count	Normalized occurrences of DM "*like*"	Chi-squared value ($X^2_{(1)}$)	*p*-value	Sig.
Female	1108	36,245	305.7	0.122	0.727	$p < 0.05$
Male	1451	46,150	314.4			

Based on the above Table 2, the chi-squared test conducted in this study yielded a chi-square value ($X^2_{(1)}$) of 0.122, with a *p*-value of 0.727. Based on this result, it can be inferred that the *p*-value ($p = 0.727$) is greater than the significance level ($\alpha = 0.05$), indicating that this present research should accept the null hypothesis (H_0). Therefore, this present research concludes that the usage of

DM "*like*" in the corpora is independent of gender. In other words, there is no significant correlation between the usage of DM "*like*" and the gender of the speakers.

Furthermore, to measure the strength of the association between the speakers' gender and the usage of the DM "*like*", this research also performed an effect size test using Cramér's V (V) measurement. The results showing that the effect size between the two variables is 0.04 (V = 0.04). According to the result, the Cramér's V value suggests a minimal effect size. These results indicate that the variables under investigation have a limited impact or influence on each other. Therefore, based on the effect size analysis, it can be concluded that the relationship between the variables is not statistically significant and does not have substantial practical significance (Kearney, 2017). Thus, it can be said that although this research revealed that males currently use DM "*like*" more frequently than females, it still supports Gabrys' (2017) claim that the production of DM "*like*" by female and male speakers is statistically non-significant.

Additionally, looking deeper into the usage of DM "*like*" by female and male speakers, this research found that all the functions of DM "*like"* were found in both female and male corpora DM "*like*", yet the frequencies of each function were varied. Based on the improved Müller's (2005) taxonomy of DM "*like*" function by Gabrys (2017), the distribution of these functions of DM "*like*" produced by female and male speakers is presented in Table 3. A more detailed description of these categorizations will be provided in the following subsection.

Table 3. The Function of DM "*like*" Presented in the Corpora

No.	Function	Raw Occurrences		Normalized Occurrences (per 10.000 words)	
		Male Corpus	Female Corpus	Male Corpus	Female Corpus
1.	Searching expression	84	89	18.2	**24.5**
2.	Marking Number/Quantity	94	68	**20.4**	18.8
3.	Introducing Example	63	59	13.6	**16.3**
4.	Introducing Explanation	161	286	34.9	**79.0**
5.	Marking lexical focus	824	499	**178.5**	137.7
6.	Discourse Restart	159	102	**34.4**	28.1
7.	Unclassified	66	5	**14.3**	1.4

As shown in Table 3, both female and male speakers utilized DM "*like*" with the same functions, although the frequencies of occurrence varied. However, the data indicate that male speakers use DM "*like*" more frequently within their utterances than female speakers. Additionally, male speakers dominate the frequency of each function. For a visual representation of the contrasting usage of DM "*like*" between the genders based on Table 3, please refer to Figure 1.

Based on Figure 1, male speakers tend to use DM "*like*" more frequently than female speakers. Additionally, Table 3 reveals that both female and male corpora demonstrate the usage of DM "*like*" to introduce new information and focus more frequently compared to other functions. In the male corpus, speakers predominantly used DM "*like*" to introduce focused information, with a frequency of 824 occurrences or 178.5 times per 10,000 words. On the other hand, the female corpus showed that female speakers produced 499 occurrences, equivalent to 137.7 times per 10,000 words. Based on this data, the male corpus exhibits the following functions in descending order of frequency: marking lexical focus, introducing explanation, discourse restarts, marking number/quantity, searching expression, and introducing examples, with normalized frequencies of 178.5, 34.9, 34.4, 20.4, 18.2, and 13.6 occurrences per 10,000 words, respectively. Additionally, there were 14.3 occurrences per 10,000 words of unclassified instances (refer to Table 3 and Figure 2 for more details).

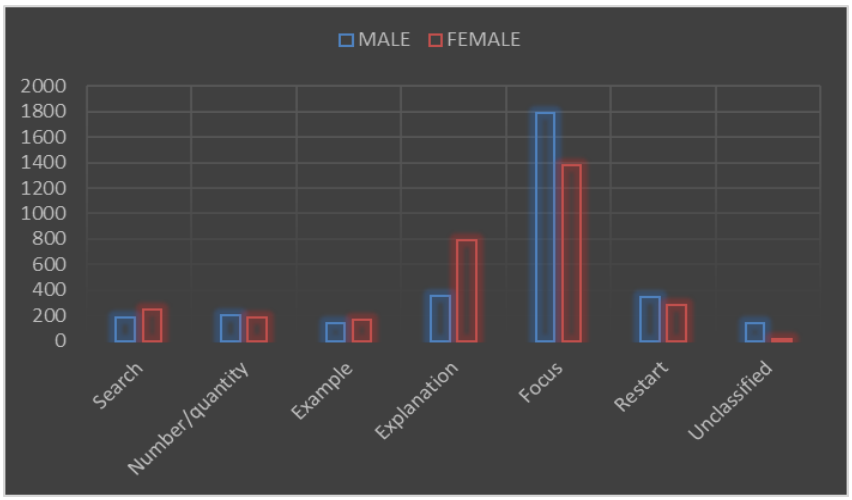

Figure 1. The Usage of Discourse Markers "*like*" between Genders.

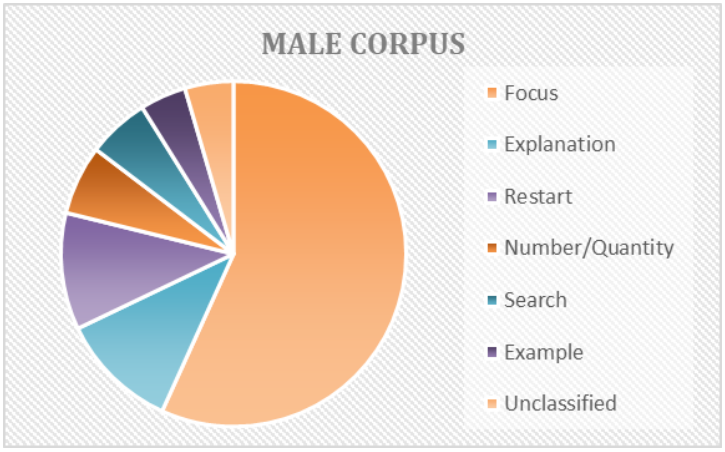

Figure 2. The Function of DM "*like*" Distribution in Male Corpus

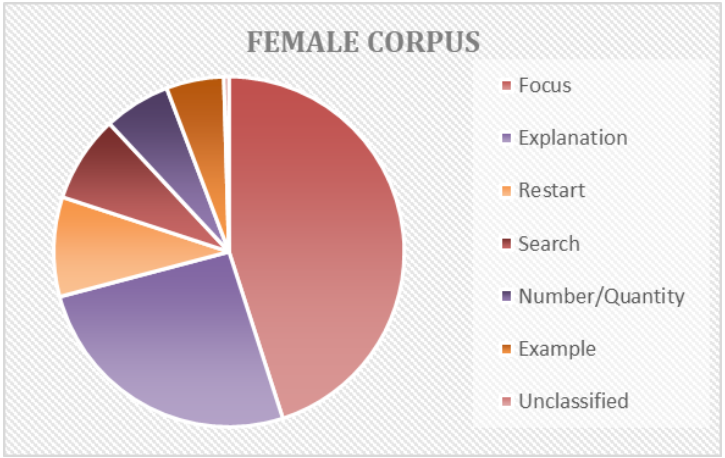

Figure 3. The Function of DM "*like*" Distribution in Female Corpus

In contrast, the female corpus showed a different frequency distribution of the functions of DM "*like*" used by female speakers. The functions, ranked from highest to the lowest frequency of occurrences per 10,000 words, were as follows: marking lexical focus, introducing explanation, discourse restart, searching expression, marking number/quantity, and introducing example. The corresponding frequencies of occurrences per 10,000 words were 137.7, 79.0, 28.1, 24.5, 18.8, and 16.3, respectively. Additionally, there were 1.4 occurrences per 10,000 words of unclassified instances. For further details, please refer to Table 3 and Figure 3.

Based on the data mentioned earlier, it can be inferred that female and male speakers employ DM "*like*" with the same functions, as explained by Müller (2005) and Gabrys (2017). However, the data indicate that male speakers dominate the usage of three functions: introducing new information and focus, discourse restart, and marking number/quantity. On the other hand, female speakers dominate the other three functions: introducing an explanation, introducing an example, and searching for expression. Furthermore, the data reveals that male speakers produce more unclassified instances than female speakers. A more comprehensive explanation of the data distribution and the differences in the characteristics of DM "*like*" usage between female and male speakers can be found in the following subchapters.

Marking Lexical Focus

Using the DM "*like*" to mark the lexical function appears to be the preferred function for both genders in this study. This assertion is supported by the high frequency of the DM "*like*" being used to mark lexical focus. Müller (2005) and Gabrys (2017) proposed that speakers employ this function to introduce new information and emphasize specific lexical items. In terms of frequency, the male corpus exhibited 824 instances of this function, equivalent to 178.5 occurrences per 10,000 words. On the other hand, the female corpus indicated that females utilized this function 499 times, corresponding to 137.7 occurrences per 10,000 words. Two examples showcasing the use of DM "*like*" to introduce new information are provided as follows. It is important for readers to note that in every analysis of DM "*like"* functions focused on the word "*like*" typed in bold based on their respective functions. Meanwhile, the functions of the other word "*like"* were mentioned in the note.

(1)

1 M: Yeah, I mean Drake dad **[like]**[1] I remember
2 Drake was **[like]**[2] playing a song on Instagram
3 live one day...
(MM-B, MC)

(2)

1 N: you know what's crazy is? **[Like]**[3] we <u>dressed</u>↓
2 [like][G] 20-year-olds. If you look at Tik Tok
3 now they all look so **[like]**[4] <u>grown</u>↓ up...
(FF-A, FC)

**Note*:

1. The superscripted numbers on each word "like" ([like][Number]) is a symbol to denote the existence of DM "*like*" mentioned in the analysis.
2. [like][G]: the word "like" as grammatical function.

As observed in examples 1 and 2, the DM "*like*" used in this function is employed to introduce new information to the listener. In example 1, the speaker intends to give a new information to the hearer regarding Drake's live Instagram activity in the past. In both examples, two DM "*like*" instances are produced in a single turn to mark a lexical focus. The difference lies in the fact that the DM "*like*" in example (1) line 1 ([like][1]) precedes new information in the form of a full clause indicated by the sentence "*I remember Drake was like playing a song on Instagram live one day*".

On the other hand, the DM "*like*" in example (1) line 2 ([like][2]) serves as a lexical focus, emphasizing the point that Drake was playing a song during the live session and expect the listener to focus on the given information.

Similarly, example 2 exhibits three occurrences of the word "*like*" in the utterance. The first "*like*" ([like][3]) is used to introduce new information in the form of a full clause, the second "*like*" ([like][G]) functions as a preposition (grammatical function), and the third "like" ([like][4]) indicates a lexical focus. In example 2 ("*we dressed like 20-year-olds*"), the speaker employs the DM "*like*" to present new information about their dressing style to the hearer during a conversation about millennials dressing style in Tik-Tok that have not been mentioned before. Conversely, the DM "*like*" in line 3 ([like][4]) marks a lexical focus and emphasizes the word "grown up," highlighting the idea that Tik-Tokers nowadays appeared to be more mature than their real age. These examples demonstrate that speakers use the DM "*like*" to introduce new information they want to highlight or emphasize within a sentence. In previous research, Müller (2005) claimed that the application of the DM "*like*" in this function includes noun phrases, verb phrases, adverbs, and adjectives but not full clauses, which he referred to as "marking lexical focus." However, the data in this study indicates that new information can be introduced in the speakers' utterances using the DM "*like*" as a full clause (example (1), [like][1] and example (2), [like][3]) when they aim to convey previously unmentioned information. These findings suggest a potential change in the DM "*like*" use among young adult speakers in 2023. This also proves that language use evolves over time.

Additionally, Müller (2005) presumed that in using this function, speakers predominantly employ particular intonation (indicated by ↓ in the example 2) and stress (indicated by underlined words in example 2) following the focused information. However, this research reveals that, in some cases, the information following the DM "*like*" simply consists of new information without any stressed words or specific intonation to indicate the speaker's emphasis or focus on that information to the hearer. Most of the time, the data indicates that stress and higher intonation are present when speakers use DM "*like*" to mark the focus. Here are two examples of DM "*like*" used to mark lexical focus followed by with a very strong emphasize (indicated by uppercase letter):

(3)

1	B:	... you're saying mean shit about them
2		**[like][5]** <u>to their face</u> in order to **[like][6]**
3		get them to want to **[like][7]** PROVE THEMSELVES↓
4		to you so it's...

(FF-B, FC)

(4)

1	B:	**[Like][8]** IT'S SO IT'S SO BAD↓ that it's [like][9]
2		at what point are you just too lost in the
3		song...

(FF-B, FC)

**Note*:
1. [like][32]: Discourse link and restart

The above example illustrates that when using the DM "*like*" as a lexical focus marker, the speaker aims to emphasize the forthcoming information, prompting the listener to focus on the given information by stressing the words (indicated by uppercase letters). As seen in examples (3) and (4), the speaker used DM "*like*" to introduce a focused piece of information.

According to Müller (2005), the use of "*like*" in the utterance is likely intended to exaggerate the expression and make it sound more dramatic. Drager (2011) explained in his study that females tend to make their utterances sound more dramatic. Given that this research indicates that males used the function of DM "*like*" to mark the lexical focus nearly twice as often as female speakers, this study suggests that male speakers can also be dramatic or hyperbolic in expressing their thoughts.

Introducing Explanation

In this category, speakers used DM "*like*" to extend the given information and precede more detailed information of what they are saying to make it understandable (Müller, 2005; Gabrys, 2017). This present study found out that both genders used this function a lot. The male corpus shows that male used this function 161 times within 180 minutes, or equal to 34.9 occurrences per 10.000 words. In contrast, the female corpus revealed a higher frequency of occurrences that is 286 occurrences within 180 minutes of podcast duration or equal to 79.0 occurrences per 10.000 words. Further, it can be concluded that female speakers use DM "*like*" to introduce explanations much more than male speakers. Below are two examples of the introducing explanation function produced by female and male speakers.

(5)

1 N: ...I remember the earrings they were **[like]**[10]
2 big earrings they were **[like]**[11] purple and
3 [like][12] big-
(FF-A, FC)

(6)

1 D: I'm [like][13] I'm laying on the bed and they
2 run and jump on it [like][14] (gesture) yeah I
3 would hate that yeah.
4 B: yeah.
5 D: That's swings for both ways that's annoying 6
on both of that.
7 B: [Like][15] they're **[like]**[16] an excited jump.
8 D: Yeah
(FF-B, FC)

**Note*:
1. [like][12,15]: marking lexical focus
2. [like][13]: discourse restart
3. [like][14]: searching for appropriate expression

These examples illustrate how speakers use the DM "like" to introduce explanations for previously mentioned topics (see example (10) ([like][10,11]), and example (11) ([like][16])). In those cases, the word "*like*" functions as a DM to clarify and provide a detailed descriptions or sometimes visual description to aid the interlocutor's understanding. For instance, in example (5), the speaker elaborates on the visual aspects of the earrings they refer to, emphasizing their size ([like][10]) and color ([like][11]). This clarification helps the interlocutor identify which specific earrings the speaker is discussing. The study also identified instances where the interlocutor employs this function in different turns to assist the speaker in explaining their ideas, as seen in example (6) ([like][16]). In summary, this study found no notable differences in the usage of this function between female and male speakers, aside from the difference in frequency. The discrepancy in frequency suggests that female speakers are more inclined to provide explicit and detailed information and illustrations on specific topics than male speakers.

Discourse link and restart

The function of DM "*like"* as a discourse link and restart, as described by Müller (2005), is used by speakers when they encounter a false start within their utterance. Thus, Gabrys (2017) suggests that using DM "*like*" indicates a self-repair following a false start or serves as a link between two parts of a sentence. Based on the data collected in this research, the male corpus reveals that male speakers produced 159 occurrences of this function, equivalent to 34.4 occurrences per 10,000 words. Conversely, the female corpus shows that female speakers only used DM "*like*" as a discourse restart

with a frequency of 102 occurrences, or 28.1 occurrences per 10,000 words. Based on these reported data, it can be inferred that male speakers produced discourse restart instances more frequently than female speakers. Take a look at the following example to determine how the speakers use DM "*like*" as a discourse restart.

(7)
1 M: Some are some are- **[like]**[17] and it it is a
2 lot of game playing because they're [like]^{Q}
3 "I'm a hot girl and I have value"...
(MM-B, MC)

(8)
1 D: it's because you know your hands show how
2 old you are yeah and she's very much-
3 **[like]**[18] I'm never really aging but I am...
(FF-B, FC)

**Note*:
1. [like]^{Q}: DM "like" as a quotation

In the provided examples, the idea of self-repair by the speakers is evident in example (7) ([like][17]), and example (8) ([like][18]). In these examples, the speakers encounter self-interruption, indicated by the dash (-) symbol, and then use DM "*like"* before restarting their utterance. According to both corpora, this study found that male and female speakers exhibit similar behavior in using this function.

Searching appropriate expression

Searching for appropriate expressions is one of the functions of DM "*like*" that speakers utilize during their thinking process, as mentioned by Müller (2005). Within this thinking process, interlocutors are likely to employ DM "*like*" to hold the floor during their turn while they search for specific terms or words, recall certain information, or simply contemplate what they are going to say next (Müller, 2005; Andersen, 2001; Gabrys, 2017). Thus, DM "*like*" functions as a terminal until the speaker finds the appropriate words. This specific function of DM "*like*" occurred 84 times or 18.2 times per 10,000 words in the male corpus and 89 times or 24.5 times per 10,000 words in the female corpus. Consequently, it can be inferred that female speakers use DM "*like*" more frequently when searching for appropriate expressions while speaking. This research result aligns with Lakoff's (1975) theory, which proposed that females tend to use more fillers compared to male speakers. According to Lakoff (1975), fillers are employed by females as a way to hedge their utterances. In the case of the DM "*like*" usage in this function, it fills pauses during the thinking process, enabling the speaker to maintain the floor (Dardjowidjojo, 2014). Below are two examples illustrating the usage of DM "*like*" extracted from the corpus. In example (10), this research found that DM "*like*" is used multiple times within a turn, serving different functions in each occurrence.

(9)
1 G: ...Nikki and I have gone through a lot with this
2 topic and I'm sure a lot of you guys have too
3 but there- has there has to be **[like:::]**[19] there
4 need- there needs to be room for people to
5 talk about this...
(FF-A, FC)

(10)

1 F: ... there's obviously some drugs that [like][20] they
2 really Teeter that line of being safe **[like][21]** (.)
3 **[like][22]** (.) a lot of these fucking opioids and [like][23]
4 painkillers is [like][24] really insane.

(MM-A, MC)

**Note*:

1. [like][20,23,24]: DMs "*like*" as marking lexical function

Example (9) was extracted from the female corpus (FC), while example (10) was extracted from the male corpus (MC). These examples demonstrate that the function of DM "*like*" within the utterances ([like][19,22,23]) is to extend the turn while the speaker thinks about what they are going to say next. In example (9) ([like][19]), the speaker used DM "*like*" to maintain the floor while searching for the appropriate word to utter. Furthermore, the usage of this function was accompanied by a prolongation in the word "*like*" (example (9)), indicating a moment of thinking. Apart from the prolongation (indicated by colons (:)), this function's usage in this example is accompanied by several repetitions of the previous or following word. As observed in example (9), the speakers actually employed the discourse marker "like" to maintain their turn until eventually, she could find the expression that she desired, namely, "there needs to be room for people to talk about this."

In some cases, the repetition occurs within DM "*like*" itself, followed by a micro pause (indicated by the period in brackets (.)) or a pause, as observed in example (10). In example (10) ([like][21,22]), the speaker used DM "*like*" to maintain the floor while searching for the word they intended to utter. The occurrence of pauses in the speaker's speech when using DM "*like*" to search for certain expressions has been explained by Müller (2005) and Schourup (1983), stating that this category is often accompanied by pausal interjections in some cases (Gabrys, 2017). According to Laserna et al. (2014), the occurrence of a pause filled with a DM is an unconscious sign of disfluency from the speaker or an automatic response to prevent the hearer from taking the turn (Laserna et al., 2014). Thus, based on the cases observed in the corpus, this study assumes that the speaker employs the usage of DM "*like*" within a pause to maintain the floor and prevent the hearer from taking over the turn while they search for the appropriate expression in their mind (Laserna et al., 2014).

In some cases, speakers also combine DM "*like*" with other DMs, such as "*I mean*" and "*you know*," or certain gestures or expressions that indicate the speaker is trying to find an appropriate word while using this function. However, there are no distinguishing characteristics between female and male speakers in the usage of this function. In other words, female and male speakers use this function in a similar manner but differ in frequency.

Furthermore, this research discovered some intriguing instances where speakers used DM "*like*" in the final position, indicating a loss of words to explain what is in their minds. In these cases, the speakers were not actively seeking the appropriate expression but somewhat halted their turn. Occasionally, the speakers would utter DM "*like*" followed by a pause or accompanied by a particular facial expression or body gesture that conveyed their lack of ideas for what to say next. While this occurrence was not prevalent in the corpus, it was noticeable in both the female and male corpora. Take a look at the examples below to observe how speakers utilize DM "*like*" when they are unsure of what to say next.

(11)

1 B: ...think what if they run and jump on a bed
2 [like][G] it looks comfy yeah you know **[like][25]**
3 (gesture). You know what I mean.

(FF-B, FC)

(12)
1 M: ...what are we doing for lunch and I'm [like]^{Q}
2 "oh you know we could-" and then I'm just trying
3 to **[like]**^{26} (.) you know.
(MM-A, MC)

**Note*:
1. [like]^{G}: the word "like" as grammatical function
2. [like]^{Q}: the word "like" as quotative

The examples mentioned above demonstrate how speakers use DM "*like*" to indicate that they have lost a word and are unsure of what to say next or are unable to articulate it adequately. Moreover, in order to ensure that the listener comprehends their intended meaning, speakers may employ additional DM such as *"you know what I mean"* in example (11) and "*you know*," as illustrated in example (12), implying that they expect the listener to understand without requiring further explanation. Unlike the usage of DM "*like*" to search for suitable expressions, where speakers are eager to continue their utterance, the use of DM "*like*" in instances where speakers have lost a word often results in the termination of their turn without making an effort to find the appropriate expression to continue their statement.

Marking an approximate Number/Quantity

The function of DM "*like*" in introducing specific numbers, such as weight, height, age, dates, and frequency, primarily serves to emphasize the numerical expression that the speaker intends to convey to the listener. Additionally, in some instances, the use of DM "*like*" before mentioning a number indicates that the speaker is somewhat uncertain about the exactness of the statement (Müller, 2005; Gabrys, 2017). Apart from preceding the number, DM "*like*" can also be employed to introduce quantity, time periods, or frequency (Müller, 2005). When considering the data presented in this study, it becomes apparent that males are more likely to use this function than females. In the male corpus, this function appeared in 94 occurrences, or 20.4 occurrences per 10,000 words, while in the female corpus, it occurred in 68 instances or 18.8 occurrences per 10,000 words. Hence, it can be inferred that males use this function more than female speakers. The following examples illustrate the usage of this function:

(13)
1 B: ... I'll sit in my car for **[like]**^{27} <u>30 minutes</u>
2 because I don't want to go in...
(FF-B, FC)

(14)
1 N: Um um I'd say **[like]**^{28} early in my adulthood
2 maybe **[like]**^{29} my early 20s that was
3 definitely that was...
(FF-A, FC)

(15)
1 A: ...dishware from my house but I have no
2 travel cups **[like]**^{30} <u>NONE</u> at all ...
(FF-B, FC)

In example (13) ([like]^{27}), the usage of DM "*like*" is due to the speaker's uncertainty about the exactness of the number mentioned. The speaker indicates that they sat in the car for approximately 30 minutes because they did not want to enter the building. In this case, the speaker cannot be sure about the precise duration of time. Example (14) demonstrates the use of DM "*like*"

to mark a non-numerical time period expression. According to Müller (2005), time periods mentioned by speakers are not always specified as numbers but can also include various expressions such as "*twenties*," "*forties*," "*a long time ago*," and so on. The exact time period is not crucial; instead, the speaker's primary objective is to convey a rough idea of the value or time period they have in mind. In example (14), the speaker uses DM "*like*" to introduce the time period expressions "*early in my adulthood*" ([like][30]) and "*my early twenties*" ([like][29]). By preceding these expressions with DM "*like*," the speaker indicates that they do not wish to provide an exact timeframe. Not to be confused with the number that is preceded by the word "*like*", this research spotted an interesting case where the speaker preceded a number by the word "*like*" yet the word "*like*" in the utterance did not serve as a DM. Take a look back at example (2) ("*we dressed* like *20-year-olds*"). In this case, the word "*like*" typed in bold serves as an adverb, fulfilling a grammatical function. The word "*like*" in the above example has the same meaning as "similar to." Therefore, based on the context, the sentence clearly indicates that the speaker dressed similarly to 20-year-olds. Conversely, the word "*like*" as a marker of approximate number and quantity always implies an approximation. Furthermore, the word "*like*" in the above example cannot be omitted as it is an integral part of the grammatical structure. On the other hand, the word "*like*" can be considered a discourse marker (DM) if its omission does not affect the grammatical structure (Müller, 2005; Gabrys, 2017). Hence, although an utterance may contain a number preceded by the word "*like*", it does not always imply approximation of the number uttered. The analysis should be based on the specific contexts of each utterance.

Additionally, the research findings indicate that DM "*like*" is also used to mark an approximate number or quantity, which does not necessarily imply uncertainty. Female speakers, in particular, use this function to emphasize the expression of zero frequency, as illustrated in example (15) ([like][30]) with the stressed word "*like*" (underlined). However, the male corpus of this study does not contain any examples of such use. Apart from this distinction, there are no other characteristics indicating a different usage of this function between female and male speakers.

Introducing Example

The function DM "*like*" to introduce examples is primarily employed by speakers to provide illustrations or exemplify concepts during conversations. This role, according to Müller (2005), entails presenting generic concepts or descriptions, and the examples chosen by the speaker help to explain the notion being conveyed.

This function occurred 63 times in the male corpus, while in the female corpus, it occurs 59 times, suggesting that both genders utilize this function similarly. However, when considering the corpus size, the normalized data reveals that female speakers employ this function more frequently, with a frequency of 16.3 occurrences per 10,000 words compared to 13.6 occurrences per 10,000 words in the male corpus. Nevertheless, this study does not identify any distinct characteristics female or male speakers exhibited when performing this function.

(16)

1	B:	...yeah and [like][31] expressing yourself
2		which is also a form of self-care yeah so
3		**[like][32]** getting new piercings getting
4		tattoos dyeing your hair.
5	A:	the nail the lash yeah.
6	B:	**[like][33]** nails and lashes. I've told Dayson
7		forever that.

(FF-B, FC)

(17)
1 F: ...everything and just updated that game
2 **[like][34]** Counter-Strike or League of Legends or
3 all these other games...
(MM-A, MC)

Note:
[like][31]: DM "*like"* as marking lexical function

As seen in examples (16) and (17), there is no apparent difference between female and male speakers in performing this function. Both genders use this function to introduce examples that illustrate the ideas they want to convey (Gabrys, 2017). The difference lies primarily in the frequency of usage, depending on how the speaker wants to emphasize the examples within their utterances. In example (16) ([like][32]), the female speaker mentions various examples of self-expression, such as getting new piercings, tattoos, dyeing hair, and getting nails and lashes. This is performed to convey the concept of self-expression in her head to the listener. Similarly, in example (17) ([like][34]), the male speaker mentions Counter-Strike and League of Legends as examples of games he refers to. These examples indicate that the use of the DM "*like*" in this function is simply to introduce examples.

Unclassified

Regarding the last taxonomy of the DM "*like*" function, there are instances in which the speakers' usage of DM "*like*" cannot be classified into specific categories. Gabrys (2017) introduced the "unclassified" function to assist researchers in classifying ambiguous examples lacking context. Based on the data, 66 occurrences of unclassified instances were produced by males or 14.3 occurrences per 10.000 words in the male corpus. Meanwhile, the female corpus shows only five occurrences of unclassified instances or equal to 1.4 occurrences per 10.000 words that appeared in the corpus. According to this, it can be inferred that male speakers produced more unclassified instances than female speakers. Two examples of unclassified instances within the corpus are extracted as follows.

(18)
1 D: ...they call it [like][35] Halloween night or
2 something.
3 B: yeah yeah so they had [like][36] some variation of
4 [[**like**][37]].
5 D: [it was [like][38] a carnival].
6 B: yeah but everybody wore costumes
(FF-B, FC)

*Note:
1. [like][35,36,38,]: Marking lexical focus

As observed in examples (18), there are five occurrences of the discourse marker "like." The discourse markers "like" numbered 35, 36, and 38 served as marking lexical functions. However, the discourse marker "like" in line 4, numbered 37, is unclassified as the speaker did not continue her utterance. In such cases, the function of the discourse marker "like" cannot be classified since the researchers are unable to identify the speaker's intention in using the discourse marker "like" within an unfinished utterance. According to Gabrys (2017), the occurrence of such usage of the discourse marker "like" is considered ambiguous and unclassified. According to Crible and Cuenca (2017), this occurrence is common in natural and unplanned speech where the second argument following the DM is missing or incomplete, thereby blurring the DM's function. Often, the speakers do not continue their utterances due to overlapping speech from the interlocutor. Consequently, the speakers

focus on answering or responding to the interlocutor's utterance without continuing their intended message from the previous turn. The distinction between the unclassified category and the lost word category lies in the cause of the speaker's incomplete utterance. In the lost word category, speakers pause because they are unable to find the appropriate word or expression to continue their speech, whereas, in the unclassified category, speakers are compelled to halt their speech due to overlapping by the interlocutor.

Interestingly, this study did not identify any unclassified instances in the female podcast group A. Additionally, the occurrences of unclassified instances are significantly lower compared to the male corpus. Therefore, this study suggests that the lesser occurrence of unclassified instances in the female corpus indicates that females tend to complete their statements and provide clear explanations within their utterances. In contrast, based on the data, male speakers tend to quickly shift their focus and truncate their utterances during a conversation.

According to the findings, this research provides a comprehensive evaluation of the functions of the DM "*like*" and the distinct usage characteristics by each gender, as determined through the analysis of podcast data. Through the identification of specific functions attributed to DM "*like*," this study enriches our comprehension of how DM "*like*" is employed in oral interactions, particularly in terms of the different tendencies exhibited by different genders.

CONCLUSION

Discourse marker (DM) "*like"* is a common element produced by speakers during oral interactions. It not only fills pauses and maintains the floor but also highlights and emphasizes specific information within utterances or seeks the listener's attention. Furthermore, the DM "*like*" can be used to precede explanations and examples, restart the speaker's utterance, or indicate an approximate number or quantity. However, there are also instances where the use of the DM "*like*" is ambiguous due to overlapping between speakers.

By using random samples from 4 American Spotify podcasts with male-male and female-female podcasters aged from 19-40 years old, this research examined how DM "*like"* is employed in those podcasts and uncovered the different usage of DM "*like* shown by female and male speakers. Based on data collected from recorded Spotify podcast videos, this study revealed that, in 2023, male speakers produced DM "*like"* more frequently than female speakers. This finding contradicts the situation observed in the Santa Barbara Corpus of Spoken American English (SBC) from 2000 to 2005, as claimed by Gabrys (2017), where female American speakers were found to use DM "*like"* more than their male counterparts. Other previous researchers (Müller, 2005; Tolliver, 2012; Pratiwi, 2020) have also suggested that DM "*like"* is more commonly used by female speakers. However, this study brings attention to a new phenomenon where male speakers dominate the usage of DM "*like"* in podcasts which are regarded as casual face-to-face interactions.

However, despite the dominance of DM "*like"* usage by male speakers in the sample, this research confirmed that the usage of DM "*like*" is not correlated to the gender of the speakers ($X^2_{(1)}$ = 0.122, *p*-value = 0.727, *V*-value = 0.014). This result supports the findings of Müller (2005) and Gabrys (2017), which suggest that DM "*like"* and gender are not significantly associated. and gender are not significantly associated. This result proved by the data which shown that male and female speaker used all function of DM "*like*" presented in the taxonomy. Furthermore, Müller (2005) proposed that the usage of DM "*like*" across genders is primarily influenced by societal trends and not linked to gender.

This research concludes that male speakers predominantly use the following four functions of *marking approximate number, marking lexical focus, discourse restart,* and *unclassified*. On the other hand, female speakers exhibit dominance in using DM "*like"* for the other three functions usage i.e., *introducing explanation, searching expression,* and *introducing example*. Interestingly, both genders demonstrate that the most frequently used DM "*like*" function in the corpus is marking lexical focus. This finding supports Gabrys' (2017) claim that American English speakers primarily use DM "*like"* to highlight their focus on specific utterances, as observed in the SBC. Additionally, based on the data, this study reveals that DM "*like*" usage in the male corpus increases when speakers attempt to provide clear explanations.

Conversely, the female corpus shows that female speakers produce more DM "*like"* when they are excited or experiencing emotional moments during the conversation. However, the topic of conversation does not appear to influence the usage of DM "*like"* within the utterances. Moreover, this research does not identify any specific characteristics exhibited by both genders in their utilization of DM "*like*". Therefore, it can be concluded that male and female speakers employ the same functions of DM "*like"* but demonstrate different tendencies.

Additionally, within the function of DM "*like"* for searching appropriate expressions, the data indicates that both genders use DM "*like"* to indicate a lost word. This function is employed when speakers cannot continue their utterance, as if they have lost their train of thought and don't know what to say next. In using this function, DM "*like"* typically occurs at the end of the sentence before the speaker pauses. This situation is often accompanied by gestures or expressions such as "*you know*," but the speakers do not feel compelled to continue their sentence and instead stop their turn. Therefore, more comprehensive and extensive research is necessary to investigate the usage of DM "*like"* in indicating lost words.

Furthermore, this research also reveals that female speakers use this function to emphasize the absence of a certain expression, as indicated by the lexical stress on words like "never" or "none." Müller (2005) previously stated that speakers use DM "*like"* to mark approximate numbers, frequencies, quantities, and periods of time, which was confirmed by Gabrys (2017) in her study. However, the case of stressed zero-frequency expressions demonstrates that no approximation is involved while uttering these words, unlike other frequency expressions. Nevertheless, this study did not find any difference in the usage of the functions related to searching appropriate expressions, introducing examples, introducing explanations, and initiating discourse restart. The varying linguistic phenomena observed in this research prove that language is dynamic and continually evolves over time (Aijmer, 2022). Although the small corpus used in this research limits the ability to draw strong conclusions, the results suggested by the data may indicate a shift in the usage of DM "*like"* by American female and male speakers over the past 17 years.

These results presented in this research offer a new perspective in gender studies, especially in the usage of DM "*like*". Furthermore, the research contributes to the existing knowledge by presenting novel findings on the usage of DM "*like*" in American podcasts. It challenges previous assumptions about gender and DM "*like*" usage and provides updated insights into the evolving linguistic patterns among male and female speakers. Additionally, this research significantly contributes to the academic understanding of DM "*like*" by providing a comprehensive evaluation of its functions and unveiling the distinct characteristics associated with its usage by different genders. By shedding light on the nuanced complexities of DM "*like*" in oral interactions, this study advances our comprehension of its pragmatic functions and to the field of sociolinguistics.

However, it should be noted that this study has a lot of limitations, which offer opportunities for future research to explore further the usage of DM "*like"* on a broader coverage. Thus, this research indicates several avenues for future investigation. This research suggests that future researchers investigate the functions of DM "*like"* among American English speakers in different conversational settings or social groups, using larger datasets to attain more significant results.

REFERENCES

Aijmer, K. (2022). 'Did you just basically steal everything?' – A study of discourse -pragmatic variation and change. *Journal of Pragmatics*, *200*, 8–23. https://doi.org/10.1016/j.pragma.2022.07.003

Andersen, G. (2001). *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Biber, D., Conrad, S., & Reppen, R. (1998). Norming frequency counts. In *Corpus Linguistics: Investigating Language Structure and Use* (Cambridge Approaches to Linguistics, pp. 263-264). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511804489.017.

Cahyanti. R. (2021). Discourse markers In A vlog video by bnay youtube channel. *Language Horizon: Journal of Language Studies*, *9*(2). Retrieved from https://ejournal.unesa.ac.id/index.php/language-horizon/article/download/42512/36548/

Clifton, A., Pappu, A., Reddy, S., Yu, Y., Karlgren, J., Carterette, B., & Jones, R. (2020). *The Spotify Podcasts Dataset*. https://pypi.org/project/langid/

Crible, L., & Cuenca, M.-J. (2017). Discourse markers in speech: Distinctive features and corpus annotation. *Dialogue & Discourse*, *8*(2), 149–166. https://doi.org/10.5087/dad.2017.207

Dardjowidjojo, S. (2014). *Psikolinguistik: Pengantar pemahaman bahasa manusia (Kedua)*. Jakarta: Yayasan Pustaka Obor Indonesia.

Dohen, M., Schwartz, J.-L., & Bailly, G. (2010). Speech and face-to-face communication-An introduction. *Speech Communication*, *6*, 477–480. https://doi.org/10.1016/j.specom.2010.02.016ï.

Drager, K. K. (2011). Sociophonetic variation and the lemma. *Journal of Phonetics*, *39*(4), 694–707. https://doi.org/10.1016/j.wocn.2011.08.005.

Fung, L., & Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied Linguistics*, *28*(3), 410–439. https://doi.org/10.1093/applin/amm030

Gabrys, M. (2017). Like as a discourse marker in different varieties of English. Retrieved from https://dial.uclouvain.be/memoire/ucl/fr/object/thesis%3A11925/datastream/PDF_01/view.

Gaither, S. E., Cohen-Goldberg, A. M., Gidney, C. L., & Maddox, K. B. (2015). Sounding black or white: Priming identity and biracial speech. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2015.00457.

Kearney, M. (2017). Cramér's V. Retrieved from https://www.researchgate.net/publication/307963787_Cramer's_V.

Lakoff, R. (1975). *Language and woman's place*. New York: Cambridge University Press.

Li, H. (2014). *Small spoken corpus and oral English classroom*. *(ICSSR 2014)*. Retrieved from https://www.atlantis-press.com/article/11840.pdf.

Laserna, C. M., Seih, Y.-T., & Pennebaker, J. W. (2014). *Um . . .* Who Like Says *You Know*. *Journal of Language and Social Psychology*, *33*(3), 328–338. https://doi.org/10.1177/0261927X14526993

Martín-Laguna, S., & Alcón-Soler, E. (2018). Development of discourse-pragmatic markers in a multilingual classroom: A mixed method research approach. *System*, *75*, 68–80. https://doi.org/10.1016/j.system.2018.03.009

Müller, S. (2005). *Discourse marker in native and non-native English discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Nurmukhamedov, U., & Sharakhimov, S. (2021). Corpus-Based vocabulary analysis of English podcasts. *RELC Journal*. https://doi.org/10.1177/0033688220979315.

Pasaribu, T.A. (2017). Male and female students' Use of textual discourse markers in writing academic essays. *Journal of Language and Literature*. Retrieved from https://e-journal.usd.ac.id/index.php/JOLL/article/view/587.

Pratiwi, L. R, Basthomi, Y., Yannuar, N., Hidayati, M., & Fuad Choiron, N. (2020). "You Know" Aa a disourse marker in English learners' spoken discourse: UM-speaks. In *PJAEE*, *17*(7), Retrieved from https://archives.palarch.nl/index.php/jae/article/view/3096.

Ratih, E., & Abidah, K. H. (2022). The use of English discourse markers across gender: A corpus-based study. *LingTera*, *9*(1), 53-62. https://doi.org/10.21831/lt.v9i1.52970.

Rundblad, G. (2015). *Normalisation in Discourse Analysis*. Retrieved from www.appliedlinguistics.org.uk.

Schiffrin, D. (1987). *Discourse markers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511611841.

Schourup, L. C. (1983). Common discourse particles in English conversation. *Working Papers in Linguistics, 28*, 1-119. Retreived from https://core.ac.uk/download/pdf/159566699.pdf.

Taguchi, N. (2002). A comparative analysis of discourse markers in English conversational registers. *Issues in Applied Linguistics*, *13*(1). https://doi.org/10.5070/l4131005052.

Tipton, E. (2014). Stratified sampling using cluster analysis: A sample selection strategy for improved generalizations from experiments. *Evaluation Review*, *37*(2), 109–139. https://doi.org/10.1177/0193841X13516324.

Tolliver, D. (2012). *Like* and African American English: Factors and implications of language shift. Yale University Linguistic Department, USA. Retrieved from https://ling.yale.edu/sites/default/files/files/alumni%20senior%20essays/TolliverSeniorEssay.pdf.

Turhaan, N.S. (2020). Karl Pearson's chi-square tests. *Academic Journals Vol. 15(9), pp. 575-580.* DOI: 10.5897/ERR2019.3817.

Wardhaugh, R. & Fuller, Janet M. (2015). *An Introduction to sociolinguistic* (Seventh Edition). Chichester: John Wiley & Sons, Inc.