

GENDER AND LANGUAGE VARIATION ON THE COMMENTS OF VIRAL YOUTUBE VIDEOS

¹*Jovelyn Maglangit Cantina, ²Prescilla Lacaya Alvarico

¹*Department of English, College of Arts and Sciences, Jose Rizal Memorial State University

²Department of Social Science, College of Arts and Sciences, Jose Rizal Memorial State University

DOI: [10.55559/sjahss.v1i07.36](https://doi.org/10.55559/sjahss.v1i07.36)

Received: 17.07.2022 | Accepted: 26.07.2022 | Published: 27.07.2022

Electronic reference (*Cite this article*):

Cantina, J. M., & Alvarico, P. L. (2022). GENDER AND LANGUAGE VARIATION ON THE COMMENTS OF VIRAL YOUTUBE VIDEOS. *Sprin Journal of Arts, Humanities and Social Sciences*, 1(07), 341–359. <https://doi.org/10.55559/sjahss.v1i07.36>

Copyright Notice:

© 2022 Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0 : <https://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

ABSTRACT

This study aims at analyzing the language variations between female and male comments on YouTube viral videos as to abbreviations, emojis, laughter variants, and spelling variants of personal pronouns, utilizing a mixed-method design. This study revealed that males tend to use abbreviations in their comments in the leisure domain. Females used more abbreviations in the information and knowledge domains and emojis in the two domains. The female users used *haha*, *hehe*, and *jaja* more frequently than male users in the leisure domain. Male and female users used the laughter variant more often in the leisure domain than in the information and knowledge domain. Women preferred to write the standard spelling of the personal pronouns "I" and "you." Moreover, both men and women used abbreviations to express their views immediately to speed up the typing of messages. Women were more familiar with positive and negative emojis than men. Language varies according to YouTube users' preferences in using the language when posting comments online, and the core social attributes influencing language use are social class, social networks, sex and gender, ethnicity, and age. Thus, infographic material with meanings and examples can be distributed to students and teachers.

Keywords: gender variation, language variation, viral YouTube videos, comments, social media

1. Introduction

Language and technology played significant roles in creating new form of expressions, which have given rise to new words such as 'selfie' and 'swag'. Social media has made it easier to bring new words into the lexicon. A person can post descriptions that would describe

his or her actions, situations, and emotions, which could instantly start a trend that can be picked up by people around the globe. Before you know it, these words or phrases have become acceptable both in spoken and written communication and have been used worldwide.

YouTube: Broadcast Yourself™ ("YouTube") is a public video sharing Web site. The site consists of user-generated content and ranks among the most successful Web 2.0 projects. Registered users can upload their videos to the server, share them with the world, watch all other videos uploaded to the site, and interact with the community (Ammari, 2019). Nowadays, there have been a lot of videos on YouTube that have become viral, and users around the globe have posted their comments to express their views regarding these videos. YouTube personality Kevin Nalty (known as Nalts) recalls in his blog: "A few years ago, a video could be considered 'viral' if it hit a million views," but says as of 2011, only "if it gets more than 5 million views in a 3–7 day period" can it be considered "viral" (Cox, 2015). YouTube video comments are an asynchronous type of computer-mediated communication that allows the users of the Web site to respond to the viewed videos. These linguistic choices of males and females are very significant in this regard, as they reveal the overall structure of a speech community. Therefore, since YouTube is a media sharing network that is used worldwide, it is very important to study the kind of language that is being used there. The comments serve from their very nature as a subjective way of evaluating videos, evaluating other users' comments (Weder 2008) and as a way to post thoughts and opinions inspired by the video or the comments of other users.

In the field of sociolinguistics and language education, one of the key subfields is the investigation of language variation and style shifting (Jaspers, 2010 as cited in Goodman (n.d.). Variationists take the point of view that groups of speakers may exhibit unique phonological, lexical, or grammatical features. These patterns of mixing languages are shown to be systemic, and rule based. Gender-based written speech differences have attracted the attention of sociolinguists as they reveal the speech norms of particular speech communities. Several previous studies have revealed that gender influences both genders' linguistic choices (Xia, 2013). Specifically, females tend to be more willing to use color words borrowed from French in their descriptions. Research shows that although emojis are used by both males and females, females tend to use emoticons to express "solidarity, support, assertion of positive feelings, and thanks," whereas males tend not to use emojis for these reasons (Wolf, 2000).

In terms of the use of swearing terms by males and females, there are obvious gender differences. Lakoff (1975), as cited in Ammari (2019), suggests that females are always expected to make use of euphemisms (rather than swear words), while, on the other hand, males can generally say whatever is on their mind. Shawcroft (2014) found that males tended to use chat-type language to contact a romantic partner more than phone calls or face-to-face contact. Also, females prefer to argue using chat language because the time lag provides them an opportunity to consider their responses more carefully, while males seem concerned that text messages can be misinterpreted and cause further arguments.

However, because these areas have not been widely studied, the present study seeks to fill this gap in research by investigating how language use varies by gender on media sharing networks like YouTube. YouTube video comments are an asynchronous type of computer-mediated communication that allows the users of the Web site to respond to the viewed videos. These linguistic choices of males and females are very significant in this regard, as they reveal the overall structure of a speech community. Therefore, since YouTube is a media sharing

network that is used worldwide, it is very important to study the kind of language that is being used there. This study is also very important for linguistics teachers and students as it will allow them to further understand how language is used by males and females not only as a medium of interaction but as a medium for differentiation. Furthermore, findings will determine if the comments of users and non-users of YouTube may be additional beneficial sources of information for students and teachers to interpret linguistic data as they expand their understanding of language and citizen sociolinguistics. It will hopefully encourage teachers not only to view the videos themselves, but also to explore whether these corpora could enhance their own classroom practice by utilizing them as authentic data for linguistic research purposes. The significant findings of the study will be the basis for the development of infographic material which will contain information on how to be careful with language in posting comments online. This material adheres to one of the objectives of the Research and Development Office of this University, which is to produce research that would lead to cultural, societal, political, and economic transformation. The researchers will ensure utilization of research output in the community.

2. Objectives

This study aims at analyzing the language variations between female and male comments on YouTube viral videos as to abbreviations, emojis, laughter variants, and spelling variants. Specifically, it answers the following specific questions:

1. What is the frequency and classification of proportional occurrence of abbreviations, emojis, laughter variants, and spelling variants of personal pronouns, I and you?
2. Which gender is more likely to use abbreviations, emojis, laughter variants, and spelling variants of personal pronouns, I and you?
3. Which gender is more likely to favor formal or informal language registers?
4. What output can be developed based on the significant findings of the study?

3. Scientific Basis/Theoretical Framework

This study is anchored on the sociolinguistic theory of William James as cited in Chambers (2017). Sociolinguistic theory provides a dynamic view in which change is apprehended in progress. It is where 'language', or 'the linguistic', meaningfully interacts with 'society', or 'the social'. According to James, the core social attributes affecting language use are social class, social networks, sex and gender, ethnicity, and age.

Sociolinguistics is concerned with how language use interacts with, or is affected by, social factors such as gender, ethnicity, age, or social class, for instance. As Coulmas (2013) defines it, it is the study of choice and "the principal task of sociolinguistics is to uncover, describe, and interpret the socially motivated" choices an individual makes. This theory is suitable to support this study because it is a strategy that draws on YouTube users' existing knowledge, beliefs, and skills and allows them to construct and symbolize the more elaborate meanings of complex expressions. This approach provides them with a variety of ways to present their views structurally and manipulate words and phrases in order to express ideas more clearly. Sociolinguists are interested in how we speak differently in varying social

contexts and how we may also use specific functions of language to convey social meaning or aspects of our identity. Sociolinguistics teaches us about real-life attitudes and social situations.

4. Methodology

This study employed a mixed-method design. A mixed method design involves the combination or integration of qualitative and quantitative research in the research study (Creswell, 2014). The data was treated using frequency and percentage to identify the frequency of the analyzed phenomena in the corpus as a whole and in its two domains for quantitative statistical methods. A text analysis was performed for the qualitative method. So that there aren't too many differences, the corpus only includes the comments that YouTube users have left on the four most popular videos.

To obtain a representative corpus in which the identification of the most frequent formal features would be possible, this paper adopted the methodology as used in the study of Varga (2009), who carried out a thorough analysis of YouTube Video Comments. The YouTube categories were cars and vehicles, comedy, education, entertainment, film and animation, gaming, how to and style, music, news and politics, non-profits and activism, people and blogs, pets and animals, science and technology, sports, and travel and events (YouTube 2009h). These categories are listed into two domains, which contain material of a similar nature. These are the Leisure Domain (cars and vehicles; comedy; entertainment; film and animation; gaming; music; people and blogs; pets and animals; sport; and travel and events); and the Information and Knowledge Domain (education, how to, and style; science and technology; news and politics; and non-profits and activism). Obviously, both domains, regardless of their actual sizes, must be represented in the corpus. Thus, there were two videos from the L domain and two from the IK domain. The second criterion for the selection of videos is that these videos became viral in 2019-2020 and had at least five million views. Then the researchers drew and picked two categories for the leisure domain and the information and knowledge domain. After determining the categories, the researchers browsed YouTube and searched for viral videos that qualified the criteria. Figure one presents the domain, categories, and video titles in the study.

Table 1
Selected Videos

Domain	Category	Video Title
Leisure	Film and animation	The Lion King Official Teaser Trailer
Leisure	Music	Con Calma
Information & Knowledge	News and politics	The Coronavirus Explained & What You Should Do
Information & Knowledge	How to & Style	This Is The Most Expensive Plane Ticket In The World

After determining the videos, approximately 100 comment-long stretches of each of them were abstracted. All the stretches are ordered from the oldest comments, i.e., from the very first comment on the video up to the 100th one. Only comments written in English were utilized as a corpus in the study. Comments were categorized as female and male. When a user

provides his or her name to a Google account or to a Google partner, Google can determine a user's gender and age. Also, many visitors use a Google account where they provide information like gender and date of birth. But this information is not 100% accurate. The researcher went through their profiles to see if the user was female or male before determining the gender. In cases of uncertainty, the researchers communicated to the users through email to confirm their gender as either male or female. Then four groups of basic features of internet language were present in the study: abbreviations, emojis, laughter variants, and spelling variants of personal pronouns I and you were analyzed. These features were studied using mainly a quantitative statistical method to identify the frequency of the analyzed phenomena in the corpus as a whole and in its two domains (Hastrdlová 2006: 48). Furthermore, the proportional occurrence of the phenomenon in the two domains to compensate for the disproportion in size between them was provided. Next, a comparison of the proportional occurrences of the phenomena in the two domains was conducted by counting the rate. On the other hand, an analytical procedure of making comparisons in which the qualitative description of the phenomena was provided was also provided, and the phenomena were classified according to their classifications. To analyze the variations in language, the data was divided into two registers: formal and informal.

5. Results and Discussion

5.1 Basic Features of Internet Language

This section presents the frequency of the common abbreviations, emojis, laughter variants, and spelling variants of personal pronouns I and you in the corpus, describes their usage, and presents their internal classification. Also, it looks into the distribution of these features in the two domains under investigation and looks for the tendency towards the difference between them.

Table 2 lists the most frequent abbreviations and their classification as they were found in the corpus. Overall frequency is the number of times a word appears in the corpus as a whole, no matter what domain it is in.

Table 2
Frequency of Abbreviations in Two Domains

Classification	Acronym	Meaning	Overall Frequency	%	Leisure		%	Info. & Knowledge	
					Male	Female		Male	Female
Exclamation	OMG	Oh My God	25	8%	2	3	92%	4	16
Laughter Variant	LOL	Laugh (ing) out loud	19	73.68%	11	3	26.32%	4	1
	LMAO	Laughing my ass out	5	80%	4	0	20%	1	0

Swear Word	WTF	What the Fuck	6	33.3 3%	2	0	66.67%	4	0
Neutral Acronyms	COVID	Coronavirus disease	14	0%	0	0	100%	4	10
	IDK	I don't know	8	25%	1	1	75%	2	4
	TV	Television	7	71.4 2%	2	3	28.58%	0	2
	VR	Virtual Reality	7	100 %	2	5	0%	0	0
	CGI	Computer Generated Imagery	6	0%	0	0	100%	2	4
	NYC	New York City	5	20%	1	0	80%	0	4
	mRNA	Messenger Ribonucleic acid	4	0%	0	0	100%	0	4
	USA	United States of America	4	25%	1	0	75%	0	4
	3D	Three Dimensional	4	100 %	4	0	0%	0	0
	CDC	Center on Disease Control	4	0%	0	0	100%	4	0
	BBC	British Broadcasting Corporation	3	33.3 3%	1	0	66.67%	0	2
	MR	Mitral Regurgitation	3	0%	0	0	100%	3	0
	DIY	Do It Yourself	3	0%	0	0	100%	3	0
	NB	Nota Bene	2	0%	0	0	100%	0	2
	IMAX	Image Maximum	2	100 %	2	0	0%	0	0
Total			131		33	15		38	45

Note: The percentages in the columns "Leisure" and "Info & Know" show how often a certain acronym is used in each column.

Acronyms and initialisms are often found in online settings, especially in the comments of users on a YouTube video, because with them, users can type quickly. As reflected in the table, the most frequent abbreviation in the corpus is the interjection OMG, followed by LOL,

and COVID. OMG can be found both in the leisure and information and knowledge domains. However, OMG is frequently found in the video "This Is The Most Expensive Plane Ticket In The World," while LOL can be found most frequently in the video "Con Calma" by Daddy Yankee. As it can be noticed, the initialism OMG is usually used either at the beginning or end of the sentence. According to the Cambridge dictionary (2021), OMG is an abbreviation for Oh my God. OMG is used in the following utterances listed below as an interjection when someone is surprised or excited about something. In Comments 1-3, the acronym OMG can be found at the beginning, while in Comments 4-5, OMG is found at the end.

- (1) OMG the theme song is really quite same with shaolin soccer. Love it and brings all the memories back.
- (2) Omg this is crazy; I wonder how much it costs!
- (3) OMG! I have chills! I am so excited!
- (4) The thing that surprised me the most is the SHOWER! OMG!
- (5) This is so amazing omg!!!!

In a study conducted by Varga (2019), there was no instance of OMG functioning in any other way than as an ordinary interjection abbreviated for the sake of saving keystrokes. Indeed, in all the utterances listed above, OMG works as an emotional element added to the sentence (or as a holophrastic expression). There was no indication of it being nominalized or verbalized either.

LOL is the second most frequent abbreviation in the corpus. According to Crystal (2006) as cited in Varga (2009, p. 29), no one can actually tell whether the speaker really laughs when using the acronym, since no study has been conducted so far on whether people do "laugh out loud" when they post LOL. However, in the following posts listed below, we can see that the use of LOL is more or less meant to be the literal meaning of "laughing out loud."

- (6) Don't understand any word but enjoy. LOL
- (7) Did not understand a single word. Lol. but it sounds great...
- (8) Those shoes lol

The table further revealed that there were four classifications of acronyms, namely: laughter variation (LOL and LMAO); exclamation (OMG); neutral acronyms (IDK, covid, mRNA, TV, CDC, BBC, USA, NB, NYC, CGI, MR, IMAX, 3D, and DIY); and swear words (WTF).

A new way to communicate was required that would be faster to type and could also be manipulated to include feelings in the words. The new language is NetSpeak. This language uses a combination of abbreviations (including words created using digits as well as alphabetic characters), acronyms, and emoticons. Abbreviations and acronyms are used to speed up the typing of messages (Marked by Teachers, 2021).

Table 3 presents the classification and frequency of emojis in the corpus. Overall frequency is the number of times a word appears in the corpus as a whole, no matter what domain it is in.

Table 3
Frequency and Classification of Emojis in Two Domains

Classification	Type	Frequency	Over-all frequency	%	Leisure Domain		%	Info. & Knowledge Domain	
					Male	Female		Male	Female
Positive	Smiling 	103	136	97.08%	36	64	2.92%	2	1
	joking 	33		81.80%	4	23	18.90%	0	6
Negative	sad 	32	42	84.37%	6	21	15.63%	2	3
	confused 	6		0%	0	0	100%	2	4
	annoyed 	2		50%	0	1	50%	1	0
	shocked 	2		100%	2	0	0%	0	0
									
TOTAL			178						

Note: Positive emojis are pictures that express positive emotions and facial expressions, while negative emojis are pictures that express negative emotions and facial expressions.

The Internet has greatly changed the way we communicate. Since body language and verbal tone do not translate in our text messages or e-mails, we've developed alternate ways to convey nuanced meaning. The most prominent change to our online style has been the addition of two new-age hieroglyphic languages: emoticons and emoji.

Emoji are a slightly more recent invention (from the Japanese e, "picture," and emoji, "character"). **Not to be confused with their predecessors, emoji are pictographs of faces, objects, and symbols.** You're probably familiar with the distinct style of Apple's emoji: yellow cartoony faces with various expressions, as well as families, buildings, animals, food objects, mathematical symbols, and more (Grannan, n.d.).

As reflected in the table, there were 136 positive and 42 negative emojis found in the corpus, for a total of 178 emojis. The most frequent positive emoji in the corpus is the positive emoji smiling, while the most frequent negative emoji is sad.

Many studies of emoticons have found the "smiley face" to be the most widely used emoticon (Skovholt et al., 2014; Park et al., 2013, and Flesch, 2016). This study supported the existing studies. On the other hand, Drevitch (2020) asserted that the main reason for emoji use in text messages is to add more emotion to a message. The "loudly crying face" emoji can communicate both intense sadness and overwhelming joy (e.g., *I am so happy, I must cry*). In this study, the "loudly crying face" emoji means "I must cry" as illustrated in comment (11).

Furthermore, the leisure domain has the highest number of positive and negative emojis. These findings can be attributed in part to the nature of videos in the leisure domain, which are typically humorous and entertaining, whereas videos in the information and know-how domain are typically serious in nature. It can be noticed also that the users gave comments

with negative emojis in the info & knowledge domain. Well, this was attributed to their reactions to the video on COVID-19. The massive effects brought by the pandemic on our lives are really a serious matter to look into. Moreover, there was some fake news that was published, which some people might get confused about what to do and not do during a pandemic.

In this study, positive emoji come in the form of pictures that express smiling and joking, while negative emoji come in the form of pictures that express the state of being sad, bored, annoyed, confused, embarrassed, and shocked/surprised. These two classifications of emojis are illustrated in these comments:

(9) I remember watching the BBC news report on New Year's Day of 2020. I couldn't sleep, and I saw the report of a new virus. A little over a week, and it was spreading 😞

(10) I wish there was a window in the bedroom it's so nice 😊 it would just make it a lot more relaxing with the view

(11) I watched mufasa's death and I didn't even cry, it was a heartbreaking scene 😭

(12) It's funny how he sleeps and showers with his sunglasses on the whole time 😂

(13) Hope the next disney remake will be Pocahontas 😊 Lion King and Pocahontaa

(14) OMGGG I LOVE THIS SONGGGGG 😂 😂

According to Jiang and Ma (2020), the embedded positions of emojis also have three types. First, an emoji can seem like an additional element in a sentence's tail to amplify the moods or change the comment's intention. The emojis are often used to amplify a poster's emotional expression in comments. Second, an emoji can be added in a comment to visualize an object. Finally, emojis can be independently used to present a reviewer's moods and feelings. For instance, the emoji means "Acceptable", which is to agree with another's view, and the emoji means "Pray," or "High five," which is often used to express prayer for others. The three types of emojis mentioned by Jiang and Ma (2020) were also found in the corpus.

Table 4 presents the most frequent laughter variants as they were found in the corpus. Overall frequency is the number of times a word appears in the corpus as a whole, no matter what domain it is in.

Table 4
Frequency of Laughter Variant in Two Domains

Laughter Variant	Overall Frequency	%	Leisure Domain		%	Info & Knowledge	
			Male	Female		Male	Female
Haha	15	80%	3	9	20%	1	2
Hehe	10	70%	6	1	30%	2	1
Others (jajaja)	4	100%	1	3	0%	0	0

Note: The percentages in the columns "Leisure" and "Info & Know" show how often a certain acronym shows up in each column.

The third group of widely used internet languages is the laughter variant. The table shows that the majority of the users used the laughetr variant of haha among the other variants

hehe and jajaja. Further, more users used the laughter variant haha in the leisure domain than in the information and knowledge domain. This is probably because the videos in the leisure domain are entertaining. Though some studies (Tagliamonte & Denis (2008) and Varga (2009) have identified types of laughter variants such as acronyms, e.g., LOL, LMAO, emoticons, and emojis, this study considered only the various forms of the interjections haha and hehe as the laughter variants. According to Debaron (2011), The term *cyber-laugh* describes both an online joke and the increasingly common written representation of laughter by *haha* or *hahaha*, or their more sinister variants, *bwahahaha* and *mwahahaha*, transcriptions of the crazed laugh-into-one’s-cape of vampires and other stereotypical villains. *Haha is distinct from the acronymic LOL, which stands for 'laughing out loud,' a text and chat staple that occasionally makes it into spoken conversation.* It’s not an abbreviation but laughter itself, a bit of onomatopoeia—a word that stands for a sound—like *tsk, tsk* for the disapproving tongue click, or *pow!* and *oof!* for the comic-book punch thrown and received.

Tagliamonte & Denis (2008) found out that haha is the most productive laughter variant in their IM corpus, with thousands of instances in their million word corpus. In this study, there are only 15 haha and 10 hehe variants, which means this feature is not the most productive of all laughter variants. According to Varga (2019), haha and hehe naturally have various forms—they are often mistyped, repeated, and also accompanied by emoticons. In this study, almost all of the comments with haha, hehe, and jaja are accompanied by emojis.

(15) OMGG!! What is he doing.hahaha 😂

(16) omg haha 🤔 🤔 🤔 I love how he enjoy every single trips he takes! LOVE YOU

(17) Omg are you the guy on the movie warriors come out to play hehe 😂 😂 I remember that awesome movie!!

(18) Con calma jajaja 🤔

Table 5 presents the most frequent spelling variants as they were found in the corpus. Overall frequency is the number of times a word appears in the corpus as a whole, no matter what domain it is in.

Table 5
Frequency of Spelling Variants of Personal Pronouns I and You

Spelling Variant	Overall Frequency	%	Leisure Domain		%	Info & Knowledge Domain	
			Male	Female		Male	Female
I v i	9	22.22%	2	0	77.78%	5	2
You v u	3	66.66%	1	1	33.34%	1	0
U	1	100%	1	0	0%	0	0

Note. Only the spelling variants of personal pronouns I and You that are considered in the study.

The last formal linguistic phenomenon examined in this study is the spelling variants of personal pronouns I and you. Personal pronouns are among the most used words in speech. They are universal and appear in all kinds of texts, which makes them ideal for linguistic analysis (Varga, 2009).

As reflected in the table, the spelling variant I is frequent in the corpus. The spelling variant of I is only one (i), while You have two spelling variants (u and U). Furthermore, the Information & Knowledge domain is more likely to contain marked variants of 'I' while the Leisure domain contains the most marked variants of 'you'. The study refuted the findings of Varga (2019) in which the L domain is more likely to contain marked variants. Especially in the case of I and i, the rate is quite balanced, being slightly lower than 6:4. Quite unusually, the IK domain has a very similar tendency concerning 1st person pronoun I. The ratio differs from the L domain only by 6% on each side, which is really a rare observation when compared to the previous data. Flesch (2016) study revealed that other items that have a rather high frequency but are not acronyms were found, notably the non-standard spelling U for "you," which was used by 14% of Redditors.

Below are the sentences that contain the spelling variants of "I" and "You."

(16) i can't get this song out of my head 🎧👉👤

(17) When u realize that jungle book and lion king was directed by Happy Hogan:!

(18) OMG yesssssss thank U for putting snow in the video and spit sweet

(19) You know Spanish people are the kings of parties when u listen to their songs but don't understand anything and still dance 😂👉

5.2 Gender Preference as to the Use of Abbreviations, Emojis, Laughter Variants, Spelling Variants of Personal Pronouns I And You

This section presents the frequency of abbreviations, emojis, laughter variants, and spelling variants of personal pronouns I and you as applied to male and female users.

Table 6 shows how often each of the four features of internet language happen overall, broken down by domain and gender.

The table shows that males tend to use abbreviations in their comments in the leisure domain, whereas females use abbreviations in their comments in the information and knowledge domain. This means that both men and women used abbreviations to express their views immediately. By using abbreviations, they can make their writing fast and short. The study of Moss and Gunn (2009) revealed statistically significant differences in four of the five language elements, with females showing a statistically greater tendency than males to employ abbreviations (significant at the p 0.005 level). Furthermore, the study of Flesch (2016) revealed that the subcorpora contained about the same number of Netspeak types; males used 83 different items, and females 81. Females used slightly more tokens than men: 2,013 were found in the female corpus, and 1,866 in the male corpus. However, LOL was slightly more used by females; it appears 397 times in the female corpus and 332 times in the male corpus. About two-thirds of female Redditors used it, while 57% of males did. But in this study, females tended to use OMG more than males, while males used LOL more frequently than females.

Table 6
Overall Frequency as to Domains and Gender

Basic Features of Internet Language	Leisure Domain			Info.& Knowledge Domain			
	Frequency	%	Male	Female	%	Male	Female
Abbreviation	131	36.64%	33	15	63.36%	38	45
Emojis	178	88.20%	48	109	11.80%	7	14
Laughter variant	29	79.31%	10	13	20.69%	3	3
Spelling variant of personal pronouns I and You	13	38.46%	4	1	61.54%	6	2
	351						

Note: The percentage given in the columns Leisure and Info & Know is the proportional occurrence of a particular acronym in each.

This study supported the findings of a recent study conducted as cited in Atanasova (2016) in which the expression "OMG" is predominantly used by females, while the affirmation "yeah" is more strongly associated with men. Accordingly, women use words more emotionally. A recent study examined 15.4 million status updates made by 68,000 Facebook users and found that words describing positive emotions (e.g., "excited," "happy," "love"), social relationships (e.g., "friends," "family"), and intensive adverbs (e.g., "sooo," "sooooo," "ridiculously") were predominantly used by women. By comparison, male topics were fact-oriented and included words related to politics (e.g., "government," "tax"), sports, and competition (e.g., "football," "season," "win," "battle").

As to emojis, women tend to use emojis to express their opinions or ideas, both in leisure and in the knowledge and information domains. Research shows that although emojis are used by both males and females, females tend to use emoticons to express "solidarity, support, assertion of positive feelings, and thanks," whereas males tend not to use emojis for these reasons (Wolf, 2000).

According to Jiang and Ma (2020), as a visual language, an emoji expresses more information than a textual character and can be used to improve a dialogue's attractiveness as well as reduce the sentence's ambiguity. Specifically, embedding emoji into a conversational agent's language will make it more vivid and gain higher user engagement. However, Butterworth et al. (2019) indicated that gender influences emoji choices and perceptions. They then stated that people should consider how their emoji choices could impact the reception of their messages. Moreover, Prada et al. (2018) also found that female participants reported more often and expressed more positive attitudes towards using emoji than males. Nishimura (2015) also stated that 200 participants in a Japanese blog site established in 2004, had 121 major

categories and some 7000 sub-categories. He found that women are more inclined to send emojis compared to men.

Women generally rated the emotional valence of emojis more negatively than men did. In particular, negative and neutral emojis were perceived as more negative by women than by men. This is in line with previous studies suggesting that women generally tend to view negative facial emotions more negatively than men do. Women were more familiar with both positive and negative emojis than men—but not neutral ones. The higher familiarity with emojis shown by women might be a result of the finding that women also reported higher general emoji usage than men did. This effect was mostly driven by the fact that women used more emojis when texting with their families and friends. Taken together, the findings show that, on average, women have a greater knowledge of emojis than men do, and that they use them more often than men do. Also, women viewed the emotions signaled by negative and neutral emojis as more negative than men did (Drevitch, 2020).

As to the laughter variant, the female users make use of haha, hehe, and jaja more frequently than male users as to the leisure domain, while in the information and knowledge domain, the two users have the same frequency.

According to Herring (2003) as cited in de Gruyler (2013, p.124), males are more aggressive and participate more actively, whereas women use more smiles and have more instances of laughter. Despite some cases of gender-switching that have been reported in chat environments, the assumption in most gender and CMC research is that online gender usually matches the writer's offline gender on the ground that it is difficult to fake subtle linguistic cues and characteristics of female or male discourse for any sustained period of time.

In a study conducted by Ripoll and Lopez (2011), although women laugh more than men do, the daily frequency of laughter does not seem to differ. Laughter, in all its forms and manifestations, is an indicator of family vitality and healthy couples. Laughter is very attractive at the interpersonal level, especially for women. Men use humor much more and laughter when it comes to discussing sensitive health issues. In women, laughter would be more associated with greater social support in relationships and as a tool to cope with stress.

Lastly as to the spelling variants of personal pronouns I and You, the male users frequently use spelling variant both in leisure and information and knowledge domain. This means that women tend to write the standard spelling of the personal pronouns "I" and "You," while men tend to take the shortcut and use the spelling variants (i,U, and u). In the study conducted by Furtina, Fata, and Fitriasia (2016), the total number of errors in female students' writing was less than those in male students'. Hence, female students can be assumed to be more careful writers than male students. As mentioned by Salahshour, Sharifi, Nedasalahshour (2013), female students apply cognitive strategies more frequently than male students. They also added that female students are more aware of the significance of language learning and are better equipped to use language learning strategies. In contrast, Rua (2006, as cited in Furtina, et al. 2016, p.6) declared that female students tend to achieve higher in language learning, possibly due to several factors such as superior verbal intelligence, high aptitude, more motivation, social role model behavior, supporting communication, and the assumption of tasks requiring verbal interaction.

5.3 Gender Preference as to Language Register

This section presents the frequency of the users as male and female in their writing preferences, either formal or informal.

In linguistics, the register is defined as the way a speaker uses language differently in different circumstances. Registers are marked by a variety of specialized vocabulary and turns of phrases, colloquialisms and the use of jargon, and a difference in intonation and pace. Registers are used in all forms of communication, including written, spoken, and signed (Nordquist, 2019).

On the other hand, language **register** is the level and style of your writing. It should be appropriate for the situation you are in. The language register influences your writing's vocabulary, structure, and grammar. Three most common language registers in writing are: formal, informal, and neutral. But in this study, corpora are analyzed as formal and informal registers.

Table 7

Frequency of the Gender Preference of Users as to Formal or Informal Register

Language Register	Over all Frequency	%	Leisure Domain		%	Information & Knowledge Domain	
			Female	Male		Female	Male
Formal Register	55	63.64%	28	7	36.36%	17	3
Informal Register	345	47.83%	41	124	52.17%	22	158

Table 7 presents the frequency of the gender preference of the users as to formal and informal registers. As shown in the table, out of 400 comments posted by the users, the informal language register is more frequent than the formal language register. This means that the users' comments are more of a conversational type and regard other users as friends or they communicate without using standard words. Furthermore, the table revealed that, when compared to males, female users preferred to use formal registers in both leisure and information and knowledge domains. This means that the comments of male users are more in the informal register. This means that the common features of their comments contain slang, symbols and abbreviations, incomplete sentences, short sentences, jokes, personal opinions, jokes, etc. The use of first-person and second-person pronouns is a common feature of the users' comments in their preference for the use of informal register, both male and female.

In fact, people's word use differs as a function of their social groups and related characteristics, e.g., their gender and age (Sap, et al. 2014, Schwartz, et al. 2013). Abstract and analytical language—language features with greater persuasive power—are more commonly displayed by men than by women (Meier et al. 2020). Male-typical language seems to be more concerned with references to facts and "the big picture". Conversely, female-typical language has been described as more narrative, personal, social, and emotional; women tend to refer more to themselves and to other people (Meier et al. 2020). While these language

characteristics have previously been linked to lower persuasive power and status (Markowitz, 2019), language characteristics commonly displayed by women may also convey greater psychological closeness and authenticity (Joshi et al., 2020). This is in line with findings suggesting that women refer more often to affiliative topics on social media (Park et al. 2016). Unexpectedly, the same study also found that women who use social media speak a little bit more assertively than men who use social media.

The following comments listed below are examples of formal and informal register. Comments (20-22) are in the formal register, while comments (23-25) are in the informal register. Further comments (26-30) contain offensive language through their jokes, personal opinions, slang, and swear words.

- (20) Only the real ones know that this is a remake of Snow's 'Informer'
- (21) These beats are so addictive.
- (22) This is truly a masterpiece.
- (23) OMG yesssssss thank you for putting snow in the video and spit..... sweet
- (24) He deserves 10 billion views.
- (24) love this song
- (25) im dying yo this song fire but rn im bc idk Spanish
- (26) "none of this is fun" me: laughs in introvert
- (27) Dude i just wanna see you in Joker costume man 🤔🤔🔥 that will be dope.!!?👁👁
- (28) Year ago I watched lion king, something like 1 week after it released in my cinema. It was awesome. I miss the times when i was going to the cinema with my family every week to watch a movie. Now fuck corona...
- (29) They better include the slow-mo fight between Simba and Scar. Just saying. Iconic as hell.
- (30) The best part of this movie is that you didn't have to worry about the spoilers

The text comments on YouTube are subject to partial moderation, or self-moderation. Provided that the video owner enables posting comments and their rating, 11 YouTube only shows comments that obtain -5 points or better implicitly. The point system is simple: any registered user can click the thumbs-up or thumbs-down icon, which equals +1 or -1 point, respectively, depending on the user-assigned quality of the comment. The comments with less than -5 points are not displayed but can be viewed by changing the options in the Text Comments bar below the video window. The comments that have been marked as spam can also be viewed by clicking the Show button next to the spam message. Still, such a way of moderation cannot eliminate all abusive content, personal insults and undesirable off-topic posts. It is also worth mentioning that, with an increasing dissatisfaction of users with the nature of YouTube comments (Berens 2006), the Hide Comments option has been enabled and the videos can now be viewed with no text comments displayed at all.

The other specific function of YouTube comments is that of stimulating linguistic awareness. Even though many popular sources (such as Berens 2006, and Axtman 2002) deprecate the comments linguistically, many scholars (Tagliamonte & Denis 2008, Crystal 2006, Thurlow 2006) argue the opposite. It is clear that the issue can never be seen black-and-white and it should always be judged in relation to the individual person's linguistic

competence. As far as the people can switch from the casual language of comments to other stylistic levels of language, the act of writing "super-brief comments like, „crap!,, ,brilliant,, ,this guy sucks,, and ,OMG so funny,," (Berens 2006) not only cannot harm their linguistic skills, but can paradoxically stimulate their linguistic awareness. These people need to be aware of how to behave linguistically so that their comments are accepted as appropriate by the community. Otherwise, with their comments being too formal, people would be laughed at and made fun of, like in the popular article titled Ten YouTube Comments Translated into Standard English (authors' surnames unknown, Andy & Dave 2009).

5.4 Pedagogical Implications

The study affects both teaching and learning from a pedagogical perspective. When making comments on any social media platform, such as YouTube, language can differ based on gender. Equal opportunities were provided for men and women to share their views, ideas, and opinions on the videos they had watched as well as to pick up new vocabulary (NetSpeak). Although there are differences in how the fundamental components of internet lingo are used, language teachers must help their pupils recognize these differences as strengths rather than as signs of weakness or apathy. Additionally, teachers should urge the students to investigate whether these occurrences might improve their learning in addition to just watching the films. The feedback on the popular videos may be a useful source of information for future studies as well as real-world resources for illustrative purposes when teaching English as a second language. The infographic material on "Language Varies in Posting Online," which is the study's result, will be sent to DepEd students, teachers, instructors, and college students taking English writing courses.

6. Conclusion

The language used by YouTube viewers to leave their comments online differs depending on their preferences. Thus, this study supports William James' Sociolinguistic Theory, which contends that socioeconomic class, social networks, sex, gender, ethnicity, and age are the primary social factors influencing language use. Due to the entertainment appeal of these movies, users reacted more frequently with laughing in the leisure domain than in the information and knowledge domain. Both sexes utilize abbreviations to quickly express their opinions and to speed up message typing. Positive and negative emotions are more familiar to women than to males. Last but not least, women typically write the personal pronouns "I" and "you" using the usual spelling. The remarks made by users tend to be more conversational in nature and refer to other users as friends or they utilize non-standard language to communicate. Men's comments frequently contain slang, symbols and acronyms, short and unfinished sentences, jokes, personal viewpoints, and jokes, among other things.

7. Recommendation

This study suggests that language teachers might encourage both male and female students to post their thoughts online as a novel means of connecting with the rest of the world while utilizing the fundamental elements of internet language. The learners' brains may be implanted by the teachers with the idea that language variety in internet postings makes the language rich and lively. Despite the variances in how the fundamental components of internet language are used, language teachers must help their pupils see and value these variations rather

than viewing them as a flaw or a sign of disinterest. They will also encourage the students to investigate whether these phenomena might improve their learning, in addition to watching the films. The feedback on popular videos could be a useful source of information for future research projects and authentic resources for illustrative purposes in ESL instruction. Similar studies might be done to investigate other factors, such as country, ethnicity, and geographic location, which would give deeper and more detailed information about language variety. To ascertain whether or not female users used more appropriate words when writing, an inferential statistical test may be performed. Because it is challenging to draw generalizations, more research may be done using larger samples to determine whether male comments contain more slang, symbols, and abbreviations than female comments. Additionally, an infographic on "NetSpeak: Features of Internet Language" may be distributed to students and teachers. It contains the various abbreviations, emojis, laughter variants, spelling variants for the pronouns "I" and "you," and the preference for formal or informal register used by male and female users, as well as their meanings.

References:

- Andy and Dave** [surnames unknown] (2009). „Ten YouTube Comments Translated into Standard English“. DelSquacho. 15 March 2009
- Ammari, C.** (2019) Gender variation and social media: a sociolinguistic study of selected Jordanian facebook posts. <https://www.researchgate.net/publication/336103715>
- Atanasova, A** (2016, November 6). Gender-specific behaviors on social media and what they mean for online communications. *Social Media Today*.
<https://www.socialmediatoday.com/social-networks/gender-specific-behaviors-social-media-and-what-they-mean-online-communications>
- Axtman, Kriss** (2002). „,r u online?": The Evolving Lexicon of Wired Teens“. The Christian Science Monitor. 3 March 2009
- Berens, Brad** (2006). „What the Comments on YouTube Really Mean“. Mediavorous. 2 April 2009
- Butterworth, S.E., Giuliano, T.A., White, J., Cantu, L., and Fraser, K.C.** (2019). Sender gender influences emoji interpretation in text messages. *Frontiers in Psychology*, 10(1), p. 784.
- Chambers, J.K.** (2017). Sociolinguistic theory. <https://doi.org/10.1002/9781119072256.ch24>
- Coulmas, F.** (2013) Sociolinguistics: The Study of Speakers' Choices Cambridge: Cambridge University Press
- Cox, J.** (9 December 2015). "YouTube is making it easier to find viral videos". The Verge. Retrieved October 20, 2020.
- Crystal, David** (2006). Language and the Internet (2nd ed). Cambridge: Cambridge UP.
- Debaron, D.** (2011, February 27) Cyber-laugh: everything old is new again, haha. The Web of Language. <https://blogs.illinois.edu/view/25/45228>
- Drevitch, G.** (2020, February, 29). Women have greater knowledge of emojis than men do. New study offers insights into how emoji use differs between men and women. <https://www.psychologytoday.com/us/blog/the-asymmetric-brain/202002/women-have-greater-knowledge-emojis-men-do>
- Eckert, P.** (2000). Linguistic variation as social variation as social practice. Oxford: Blackwell.

- Flesch, M.** (2016) Acronyms and Emoticons on a Popular Web Forum: Does Gender Makes a Difference? A Corpus-Based Study of Reddit. *Humanities and Social Sciences*. 2016. fhal-02317528f. <https://hal.univ-lorraine.fr/hal-02317528/document>
- Furtina, D., Fata, I.A., and Fitriasia, D.** (2016). Grammatical errors in writing task: males vs. females. *Proceedings of the First Reciprocal Graduate Research Symposium between University Pendidikan Sultan Idris and Syiah Kuala University*, February 26-28, 2016, Tanjong Malim, Perak, Malaysia. <http://jurnal.unsyiah.ac.id/EEIC/article/viewFile/15845/11671>
- Hernandez, E., Hepper, K., Porter, J., Schwartzman, G., Shen, X., and Wu, Y.,** (2016). Emoji and gender. Online. Available at: <https://emojiandgender.wordpress.com/>
- Jiang, H. and Ma, J.** (2020). Genre-based emoji usage analysis and prediction in video comments. https://www.researchgate.net/publication/343229449_Genre-based_Emoji_Usage_Analysis_and_Prediction_in_Video_Comments
- Joshi PD, Wakslak CJ, Appel G, Huang L** (2020) Gender differences in communicative abstraction. *J Pers Soc Psychol* 118 (3): 417–435. pmid:31613120
- Langacker, R.** *Cognitive grammar: A basic introduction*. Oxford University Press.
- Markowitz DM** (2019) Putting your best pet forward: Language patterns of persuasion in online pet advertisements. *J Appl Soc Psychol* 19: 160–173.
- Meier T, Boyd RL, Mehl MR, Milek A, Pennebaker JW et al.** (2020) (Not) lost in translation: Psychological adaptation occurs during speech translation. *Soc Psychol Personal Sci* 3.
- Moss, G.A. and Gunn, R.W.** (2009), Gender differences in website production and preference aesthetics: Preliminary implications for ICT in education and beyond. https://www.researchgate.net/publication/220208317_Gender_differences_in_website_production_and_preference_aesthetics_Preliminary_implications_for_ICT_in_education_and_beyond/download
- Nishimura, Y.** (2015). A sociolinguistic analysis of emoticon usage in Japanese blogs: Variation by age, gender, and topic. *AoIR Selected Papers of Internet Research*, 5.
- Nordquist, R.** (2019). Definition and Examples of Acrolects in Language. <https://www.thoughtco.com/what-is-acrolect-1689057#:~:text=%22For%20%5BDerek%5D%20Bickerton%2C,by%20the%20least%20educated%20people>
- Park, J., Barash, V., Fink, C., & Cha, M.** (2013). Emoticon Style: Interpreting Differences in Emoticons Across Cultures. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*.
- Park G, Yaden DB, Schwartz HA, Kern ML, Eichstaedt JC et al.** (2016) Women are Warmer but No Less Assertive than Men: Gender and Language on Facebook. *PLoS One* 11 (5): e0155885. pmid:27223607
- Prada, M., Rodrigues, D.L., Garrido, M.V., Lopes, D., Cavalheiro, B., and Gaspar, R.** (2018). Motives, frequency and attitudes toward emoji and emoticon use. *Telematics and Informatics*, 35(7), pp. 1925–1934.
- Rafi, M.** (2010). Sms text analysis: Language. Gender and Current Practices. Available from http://www.tesolfrance.org/Documents/Colloque07/SMS%20Text%20Analysis%20Language%20Gender%20and%20Current%20Practice%20_1_.pdf
- Recamara, J.** (2021). An analysis on the attributed of students' blog posts. Unpublished Paper. Jose Rizal Memorial State University.
- Really Learn English (2010). Language Register. Formal, Informal, and Neutral. <https://www.really-learn-english.com/language-register.html>

- Ripoll, R.M.** and Lopez, R.U. (2011) Laughter: gender difference. 211(7):360-6.
doi: 10.1016/j.rce.2010.10.013.
- Sap M,** Park G, Eichstaedt J, Kern M, Stillwell D et al. (2014) Developing age and gender predictive lexica over social media. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 1146–1151.
- Schwartz HA,** Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM et al. (2013) Personality, gender, and age in the language of social media: the open-vocabulary approach. PLoS One 8 (9): e73791. pmid:24086296
- Skovholt, K.,** Gronning, A., & Kankaanranta, A. (2014). The Communicative Functions of Emoticons in Workplace E-Mails: :-)*. *Journal of ComputerMediated Communication*, 19, 780-797.
- Sharma, D.** and Rickford, J. (2019). AAVE/Creole copula absence A critique of the imperfect learning hypothesis. Queen Mary University of London / Stanford University. Retrieved on January 30, 2020 from <http://www.johnrickford.com/Portals/45/documents/papers/Sharma-and-Rickford-2009a-AAVE-Creole-Copula-Absence.pdf>
- Shawcroft,** Sara R., "Gender differences in text messaging" (2014). Available from <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=4964&context=etd>
- Sheehan, K.** (2015). The impact of direct writing conventions instruction on second grade writing mechanics mastery. NSUworks, Anrahan S. Fischler School of Education.
- Tagliamonte,** Sali A., Derek Denis (2008). „Linguistic Ruin? LOL! Instant Messaging and Teen Language“. *American Speech*, Vol. 83, No. 1, Spring 2008 [viewed online, original pagination not provided]. American Dialect Society.
- Weder,** Mirjam (2008). “Form and Function of Metacommunication in CMC“. In: Kelsey, Sigrid, Kirk St. Amant (eds.). *Handbook of Research on Computer Mediated Communication* (Vol. 2). Hershey: Information Science Reference, 570-586.
- Wolf,** Alecia. (2000). Emotional expression online: Gender differences in emoticon use. *CyberPsychology & Behavior*. 3. 10.1089/10949310050191809.
- Xia, X.** (2013). Gender differences in using language. *Theory and Practice in Language Studies*, 3(8). <https://doi.org/10.4304/tpls.3.8.1485-1489>
- Zhenpeng, X.,** Xuan, W., Wei, H. Huoran, Q., Qiaozhu, X. (2018) Through a gender lens: Learning usage patterns of emojis from large-scale android users. In *WWW 2018: The 2018 Web Conference*, April 23–27, 2018, Lyon, France. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3178876.3186157>
- Varga, T.** (2009). Language of youtube videos comments. Retrieved on January 29, 2020 from https://is.muni.cz/th/m3mm6/Varga_Thesis.pdf