Linguistic Deviations: Characteristic of Online Language

Alma H. Ramadhani
Department of Languages and Literatures
Universitas Gadjah Mada
almahwd23@gmail.com

ABSTRACT

This study presents the analysis of a small corpus of online language taken from microblog site Instagram, examining the language characteristic focusing on the linguistic deviations. As a representation of CMC, Instagram allows its user to do a free communication with another users talking about personal experience or discussing anything that matches their interest without any limitation of characters. The data used in this research was in the form of words, phrases, clauses and fragments taken from Instagram Posts’ comments. The result shows that linguistic deviations found in Instagram are non-standard capital letters, over-used of punctuations, emoji, repeating characters, laughter, abbreviations, and grammatical deviation. Online language is believed to be creative and unique and those linguistic features indicate that online language on is a combination between written and spoken communication and it has opened avenues of online language that is not accessible with “correct English”.

Keywords: Online Language; CMC; Linguistic Features; Linguistic Deviation

INTRODUCTION

The development of society and technologies has brought many impacts, no exception to the changing of the communication contexts, causing the emergence of new varieties of language use as a result of the computer-mediated communication technology. As linguist David Crystal ever said on his interview, the development of technologies gives a big contribution to the development of language, including the internet. People of different ages use using e-mail, instant messaging, and chat rooms to exchange written messages (Randall, 2002). CMC has developed rapidly, effecting the use of language complete with its features (grammar and vocab). David Crystal (2006) then initiated the term “Netspeak” to refer to the variety of language that is used in the CMC platforms.

The fast advancement of the internet has become the motivations for linguists to conduct researches focusing on its language (Baron, 2003; 2004; Crystal, 2006; Herrings & Danet, 2007; Tagliamonte, 2008, 2015; Tag & Seargeant, 2014; Villares, 2017; McCulloch, 2019). Language that people use on the internet has created functions that correct English might not have: whether they use perfect grammar or make everything shortened, use non-standard capital letters to emphasize the points, or use multiple exclamations and other punctuations to show what they feel. The uniqueness of Netspeak features which are artful, playful, and stylized (Danet & Herring, 2007) has made Netspeak become one of the richness of languages.

When people do a communication with each other using device-based mediation, the phenomena refers to what is known as CMC (computer-mediated-communication) (Kiesler, Siegel, & McGire, 1984). David Crystal (2006) then introduced the term “Netspeak” to replace CMC because both of the terms have different implementation: CMC puts its focus on the medium while “Netspeak” refers to the language which displays unique features. Beside “Netspeak”, another popular terms which people often use to refer to the same language are online language, internet language, and digital language. Unfortunately, unlike Crystal who regards “Netspeak” as a representation of the language richness, such creative ways have been criticized as deviant spelling even though the users might do that consciously because of some factors (ElBekraoui, 2008).

People write all the time now and what they write is mostly informal. There were difficulties in studying informal writing before the existence of the internet even though informal writings
have already there in the form of diaries, postcards, and letters. On its way, internet language is believed as one of the informal writing varieties but now, it seems like many linguists agree to give an extra attention to the development of internet linguistics and start to consider it as new variety of language. Internet linguistics is not just a study of the “what is the latest popular memes today”? It is a deeper look into day-to-day language that we have ever been able to see (McCulloch, 2019).

Internet language has been subject to academic studies and appears in journal from many fields. Since many studies regarding this issue have been conducted, relevant researches will be discussed in order to underlie this work. Baron (2004) conducted a study about linguistic features of CMC, in this case Instant Messaging, based on a corpus collected from American college students. The noteworthy finding of her research was that only 31 out of 11,718 words had abbreviations in which suggested as the typical features of instant messaging, 90 words were acronyms, and 49 were in the representation of emoticons. Baron’s findings proved that instant messaging was still conventional. Moreover, the findings also showed the use of lexical contraction which is more common in spoken language.

Tagliamonte & Denis (2008) examined a corpus involving 72 teenagers and over a million words of natural, unmonitored instant messaging. The result showed that the use of short forms, abbreviations, and emotional language was small, only about 3% of the whole data. Furthermore, the result reinforced earlier related research in showing that language used in instant messaging is part of broader contemporary trend towards more informal language. Another study regarding linguistic features of CMC was by Tio (2010) who investigated 197 e-mails and 71 memos from 10 subjects (5 males and 5 females). Some spoken linguistic features such as the pronoun and article omission, abbreviation, exclamation mark, multiple letters and punctuation for emphasizing, and also the non-standard use of capital letters were found. A pilot study of linguistic conducted by Dino & Gustilo (2015) investigated the features and functions of online language used by Filipino Facebook users. The result of the pilot study revealed that Filipino Facebookers used eight types of linguistic features which are abbreviation, acronym, discourse particles, borrowing, affixation, compounding, blending, and code-switching. Moreover, Villares (2016) conducted a relevant study using a more modern CMC platform. She examined the linguistic deviations and the non-standard features of language use in Tumblr. The result of the analysis revealed that text on Tumblr has features of the written communication combined with features of spoken discourse.

Summary of the Previous Studies

<table>
<thead>
<tr>
<th>Year</th>
<th>Researcher</th>
<th>Medium</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>Ferrara et al.</td>
<td>Instant Messaging</td>
<td>Subject pronouns omitted; Copulas deletion; Article deletion; Tursed sentences; Shortened words;</td>
</tr>
<tr>
<td>2003</td>
<td>Palfreyman &amp; Al Khalil</td>
<td>Instant Messaging</td>
<td>Use of offensive words; Use of Roman key-board set to write words in Arabic; Local dialect</td>
</tr>
<tr>
<td>2004</td>
<td>Baron</td>
<td>Instant Messaging</td>
<td>Turn taking; Abbreviations; Contractions; Emoticons;</td>
</tr>
<tr>
<td>2008</td>
<td>Tagliamonte &amp; Denis</td>
<td>Instant Messaging</td>
<td>Abbreviations and emotional language;</td>
</tr>
</tbody>
</table>
In the past decades, many linguists have examined online language and made similar claims, underlining its hybrid nature: the combination between formal and informal, and also written and spoken registers (Tagliamonte & Denis, 2008). Randall (2002) has argued that deviations of “correct” English are more noticeable in informal register and relying on that hypothesis, this study was purposed to examine the online language characteristics focused on its non-standard linguistic features.

When it comes to the register analysis, comparison is needed in order to gain effectiveness (Biber & Conrad, 2009). It is nearly impossible to know the distinctive features of a register without comparing it with other registers. The most reasonable way of comparison is by comparing CMC register with more-formal-and-conventional register.

**Comparison of formal and informal registers by Beiber & Conrad (2009)**

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Formal register</th>
<th>Informal register</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participant</strong></td>
<td>Monologue</td>
<td>Dialogue</td>
</tr>
<tr>
<td><strong>Platform</strong></td>
<td>Print</td>
<td>computer</td>
</tr>
<tr>
<td><strong>Time-dependency</strong></td>
<td>longest time to prepare</td>
<td>long time to prepare</td>
</tr>
<tr>
<td><strong>Durability</strong></td>
<td>most</td>
<td>Considerable</td>
</tr>
<tr>
<td><strong>Editing</strong></td>
<td>Considerable</td>
<td>Some</td>
</tr>
</tbody>
</table>

Linguistic analysis of a register is relying on the register features which could be words or grammatical characteristics that are frequently occur. Hence, one thing that should be underlined before doing a register analysis is knowing the differences between register feature, register marker, genre marker, and style feature because those terms are sometimes overlap. A summary below provides brief explanations about the differences of the four terms based on what has been suggested by Biber & Conrad (2009:55).

**Register Feature, Register Marker, Genre Marker and Style Feature by Bieber and Conrad (2009)**

<table>
<thead>
<tr>
<th>Type of linguistic characteristic</th>
<th>Complete text of Text samples</th>
<th>Distribution</th>
<th>Frequency of characteristic in a particular variety compared to other varieties</th>
<th>Use of the characteristic in other varieties</th>
<th>Characteristic functional or arbitrary/functional</th>
</tr>
</thead>
</table>

Lowercase letters for I and you; Personal pronouns; Intensifiers; Quotative system; Future temporal; Reference; Modals of necessity;
Language purist believes that electronic communication is corrupting the grammar of the “correct” language. Grammar in online language must be understood somewhat as linguistic deviations even though many linguists believe that it is not. The set of grammatical features of online language is referred to what is called e-grammar (Herrings, 2012) even though the term is not implying that there is a specific grammar for all varieties of online language. The set of features are including typography, orthography, morphology, and syntax. Online language grammar also differs in that, as a new and still developed phenomenon, it has not yet to have “rules” since the patterns vary according to the platform and situational contexts (Herring, 2007). The explanation below provides the summary of the features containing typography, orthography, morphology, and syntax based on what Herring (2012) has suggested.

1. **Typography**

Typography refers to the use of non-alphabetic keyboard symbols such as numbers; repeated punctuation; and special symbols such as &, @, and etc. Typography also includes the use of non-standard capitalization and the substitution of numbers or letters for words.

2. **Orthography**

Non-standard orthography includes abbreviation; phonetically-motivated letter substitution (e.g., d for t, z for s); spellings that imitate dialectical pronunciations; and spellings that contain sounds such as “heloooooo” or “heeeeiiii”.

3. **Morphology**

When it comes to morphology, it includes clipping, blending, acronyms (lol, OMG, WTF, jk), semantic shift, and conversion. These features are not typical to online language but they are productive on the internet, generating new words that are open their way into dictionaries of Standard English.

4. **Syntax**

When it deviates from standard syntax, the syntax of online language is described as telegraphic and fragmented. In informal style, articles and subject pronouns maybe elided and message that do not contain a complete grammatical clause are common in CMC. What caused sentence fragments is may be people’s behavior in typing speech-like utterances and/or the requirement in some CMC platforms that messages should not be no longer than certain characters (e.g., Twitter only provides 280 characters per tweet) which can lead users to break longer utterances into several messages (Baron, 2010).

CMC genres are evolving from time to time and one of the most popular among others nowadays is the microblogging site. As it is named microblog, the major differences between microblog and the ordinary blog the size of the contents. Microblogs allow its user to exchange and share short element of contents such as short sentences, pictures, and short videos (Kaplan & Haenlein, 2011). Instagram, as a representative of an informal register, is one of the most popular platforms with more than one billion active users per month. It is accessed mainly by teenagers.
and young adults, allows its users to do one-to-many communication and to write captions or comments without giving them maximum characters. As a newer and modern microblog platform, Instagram could be accessed through mobile application or website. This study used a small corpus of 230 comments taken form Instagram used random sampling so that all the comments were taken randomly from random posts. All the texts were taken by re-writing and compiling them into one file so that a small corpus was formed. Relying on what Herring (2002) has suggested, the advantages of using random sampling are the representativeness and generalizability. The data, which are in the form of words, phrases, clauses, fragments, and sentences was analyzed qualitatively by observing the discourse phenomena.

RESULT AND DISCUSSION

1. CAPITAL LETTERS AND PUNCTUATIONS

Non-standard capital letters and spelling are considered as typography. The use typography on the internet have been analyzed structurally as an illustration of expressivity.

(a) And I thought the drama in MY life was bad ??
   b. no one is going to remember yours!!!
   c. ME TOO AHDNFBDFS
   d. I DonT give off ViBEs lmao
   e. velma from scooby do when she popped out in that leather suit..... aksjdbdn
   f. check yo self ??

The way people use online like whether using perfect grammar, shortened everything, using capitals and extra letters or characteristics for emotion has opened avenues of online language that is not accessible with “correct English”. In online language, when people write everything in capital letters, they are believed to be trying to emphasize what they are saying. Back to 1984, caps lock could mean shouting or yelling but, nowadays generation has made caps more special by giving it more “meaning” such as indicates strong feeling. Randall (2002) has stated that online language is unique because it could create an actual tone in written media, with written linguistic features, but with conversational tone. By reading those caps, the vision on how that expression might be expressed is recorded in our mind. We can imagine people’s voice: whether they are screaming or whispering. Beside the use of capslock, another typography was the use of “yo” for “your” as in example (1:f). Non-standard spelling could also be used to investigate the dialects (McCullogh, 2019) because deleting the “r” is known as the feature of Southern American English and African American English.

Another typography features is the non-standard use of punctuation. Peterson (2011) suggested that punctuation on the internet aims to express emotion in platform-mediated communicative situation which is not possible to maintain face-to-face.

(b) a. Did you not read the story??? Her last name is clearly Cecelia, first name being Aunt.
   b. PJ FROM GOOD LUCK CHARLIE???
   c. He’s asleep I can’t wake him up!!!
   d. I work at McDonald’s and this is how it is ????
   e. idk if it’s an international thing but we say extra virgin olive oil, don’t we?
   :’) maybe a second meaning...

2. EMOTICONS

The use of basic emoticons such as :( and :) and another punctuation-emoticons representing faces and objects also occurs in the data sample, parallel with David Crystal (2006)’s idea that one of the defining typographical characteristics of online language is the use of emoticon. Emoticons function to “express user’s emotions, producing positive judgments among users but sometimes they can alter the meaning of the message or even invalidating its proportional content altogether” (Yus, 2011:106). Beside the conventional emoticons that people often use such as :(, :, T_T, ‘—’ , etc. emojis were also seen to give more colors in people’s online communication. People tend to use 😊 instead of the
Emojis have successfully changed the conventional-uncolored emoticon into modern-fancy-colorful graphics, gave people a lot more choices to express their feelings and to make their conversation more "interesting".

3. REPEATING LETTERS

Repeating letters, which categorized as non-standard orthography, also considered as one of defining online language characteristics and e-communication. It often manifests spelling behavior that suggest loosened orthographic norms (Herring, 2002). Non-standard orthography includes: (a) abbreviation (acronyms, clipping, vowel omission, etc.); (b) phonetically-motivated letter substitution (gr8 for great); and spellings with repeating letters that create sounds such as “hahahah” or “helloooo”, laughter, and other noises. Findings below show the use of repeating letters by people on Instagram.

(3) a. me toooooo, oof the hair, the boobs, the cool face paint, and her kickassness
b. lashes really take ya from bitch to biiiiiiiitch
c. I fuuuuuuuuccccckkkiiinngggg REFUSE to be unhappy in 2019, idgaf who gotta go
d. loooooooool

One thing that should be noticed is that there are other features that occur but since the size of the corpus made as sample is small, the number of the occurrence might be not as much as is expected but, the most important is such features do exist.

4. LAUGHTER

Since the sound of laughing belong to speech, people need to find a way to bring that sound online. It is amazing how a set of letters such as wkwk or a set of numbers such as 555555 creates a sound of laughter. But in English, the common laughter sound are still the hahahaha.

(4) a. haha I’m in this by myself buddy thx tho it be like that sometimes
b. I’m dead hahahah
c. if you’re watching someone be disrespectful to a customer service rep and you’re a customer, stand up for them. what are they gonna do? fire a customer? like fuck outta here haha
d. HAHAH LITERALLY

e. hahaha i do that

Another way to laugh when you are on the internet which people mostly do is by using lol, rotfl, and lmao. Interestingly, those abbreviations which stand for laughter are be able to create different sound: laughing with lol (laughing out loud) is less “laughing” than laughing with lmao (laugh my ass of). If we look deeper, it is mind-blowing how people could differ meanings from certain words that have a same purpose, and in this case the purpose is to show laughter.

ABBREVIATIONS

The most common abbreviations that are used on the internet are omg, lmao, and lol even though there are another abbreviations such as idk for “I don’t know”, bff for “best friend forever”, and vowel omission ppl for “people”. Another abbreviation that was found is idgaf which stands for “I don’t give a fuck” and wth which is “what the hell.”

(5) a. lol my life
b. gonna use this lol
c. omg this is great
d. **omg** love it ??

e. exactly, now you're getting it good job **lmao**

f. **ik lmao** I'm a good liar

g. I hate **ppl** say men are dogs... like I don't think a dog has ever made me cry or question my self-worth

5. **THE “URBAN DICTIONARY” LANGUAGE**

Urban dictionary is an online dictionary founded by Aaron Peckham in 1999. It consists of slang and cultural words or phrases that is not found in standard dictionaries. Slang, and any other “informal” words seems like typical characteristics of any informal communications whether it is written or spoken. Teen have been notorious for having unique vocabularies because they are able to create words creatively and then give them new meanings. Back to 80's, when kids started to say **gag me with a spoon**, followed by 90's popular **bling-bling** and **cha-ching** that we hear in Jessie J's song are proof that people have already had those “creative” behaviors back then. The massive use of **yolo** is another example of youth's creativity in creating such words.

(6) a. if you're watching someone be disrespectful to a customer service rep and you're a customer, stand up for them. What are they gonna do? Fire a customer? Like **fuck outta here** haha

b. honestly at this point Florida is kinda **no mans land**

c. if your pets don't like him, he **aint the one**

d. Don't have a **doggo** but i will smoosh all my cats right now

e. If she acts like a bitch you better **bang her ass** and treat her like a bitch????

On the example (6:b), people use “no mans land” to refer to a certain place. Back to World War 1, no man's land was used to describe the land between two enemies, which neither side wished to cross nor seize because of the fear of being attacked. Now people could use the phrase in every occasion even give it more meaning.

6. **GRAMMAR**

The syntax of online language deviates from “standard” syntax because sometimes it is described as 'telegraphic' and fragmented (Herring, 2002). In informal style, articles and subject pronouns may be elided and messages that do not have a complete grammatical components (subject and predicate) are common.

(7) a. awe he's gorgeous ????

b. death a long time ago

c. gonna use this lol

d. why u do this??

e. watching the labyrinth

f. no I just never heard someone say "non white people” before. It's funny lol

g. literally my entire relationship to catsmarina.odettel cried like a little bitché when I lost my cat for 10 minutes. We found him in the sewers

The frequency of grammatical function words (pronouns, determiners, modal auxiliaries) in electronic corpora have been measured by many researches. CMC is different from the traditional speech and
writing because it is more like speech but with writing linguistic features. In short, online language is spoken but written.

CONCLUSION

In a small corpus consists of 230 Instagram posts’ comments, the linguistic features which are considered as “deviation” towards “correct” English that occurred were the use of non-standard capitalization, or in another words, over capitalized and also the non-standard puncuations. When people use over capitalized online, they want to create a tone of voice that cannot be “heard” because it belongs to the speech. The use of over puncuations also give quite similar function: to create sounds. “I've told you !!!!” sounds angrier than the casual “I told you”, and “HAPPY BIRTHDAY” sounds more happy than the ordinary “happy birthday”. Another distinctive feature of online language is the use of emoticons (Crystal, 2006). Repeating letters, laughter, abbreviations, and grammatical deviations were also found. Those features strengthen many linguists’ claim towards online language which says that online language is the combination between spoken and written language. It is amazing how people could produce tone of voice which belongs to speech using writing features.

Even though there has been an international perception which believed that computers and mobile phones are affecting the everyday use of language and the effects were definitely not good (Baron, 2003), many linguists now agree that online language deserves a better place. Online language is no more linguistic “bastardization” (O’Connor, 2005) or “linguistic ruin” of the generation (Axterm, 2002). Baron (2003) argued that it is not people’s saggy attitude towards spelling and grammar but, it is the characteristic of what is called language “Whateverism”. However, Randall (2002) believed that online language is unique because it is not just simply a new way to exchange messages, pass notes, or send letters. The ability of combining writing and speaking that online language has is precious and unique. It is a creative language.

We are clearly tapping into a vibrant new medium of communication complete with its own unique characteristics (Ferrar, Brunner, and Whittemore, 1991; Crystal 2001; Randall 2002). This study’s aim to look for online language characteristics focusing on its deviations reveals that in this pop culture, with the help of technology, people’s creativity in using language increases. Language is not just a tool to communicate, it is also a tool to reflect its speakers mind. The word formation and the creation of new words are popularized by online people which are mostly teenagers in social network mainly because the existence of puns, memes, word plays, and another specific types of writing (Young, 2013).

The result of the present analysis are consistent with the theoretical approaches that describe the characteristics of online language and the claim which says that online language is “written but spoken” (Crystal, 2006; Randall, 2002) with new approaches of expression and language manipulation (Villares, 2017). Even though the small corpus shows that most of the people still use standard linguistic features and “correct” English, linguistic innovation which is typical of online language illustrates that “the type of language that is being created online is affecting day-to-day speech patterns and writing styles of most young adults” (Baheri, 2013).

The small number of data in the small corpus is representing the total population. The larger the corpus, the more vary the findings would be. The small number of data samples shown in the discussion section might be increased if the study was conducted based on a larger corpus. Since this study is only a small part of a larger study regarding to online language characteristics, I would make a larger corpus so that the findings could be more representative and generalize. Observing on how such language change is indeed interesting. How language evolves from time to time and creates new register is something that often escapes our attention even as a linguist. It is hoped that the analysis of the present study could illustrate how the digital medium gives people triggers and lead them to create a “new kind” of language, combining both spoken and written features. Further researches in the field of online linguistics still need to be done since the technology and language would not stop evolving.
REFERENCES


Baheri, T. S. 2013. Your ability to can even: A defence of Internet linguistics. The Toast.


