Investigating Indonesian language varieties in social media interactions: Implications to teaching practices

Hari Bakti Mardikantoro[1*], Muhammad Badrus Siroj[2], Esti Sudi Utami[3], and Endang Kurniati[4]

[1]*Indonesian Literature Study Program, Faculty of Language and Art, Universitas Negeri Semarang, Sekaran, Gunungpati, Semarang, Indonesia*
[2]*Indonesian Language and Literature Education Study Program, Faculty of Language and Art, Universitas Negeri Semarang, Sekaran, Gunungpati, Semarang, Indonesia*
[3,4]*Javanese Language and Literature Education Study Program, Faculty of Language and Art, Universitas Negeri Semarang, Sekaran, Gunungpati, Semarang, Indonesia*

ABSTRACT

Social media users come from various circles with different social and cultural statuses which will bring up diverse language variations. Additionally, language variations in speech events on social media certainly have characteristics that are distinctive and different from speech events in general. This study aimed at investigating types and forms of Indonesian language varieties occurred in speech events on social media. A quantitative and qualitative approaches was applied to accomplish the research objective. The quantitative approach was carried out using content analysis. The qualitative approach used was based on sociolinguistic theory. The research data were collected using the listening method (uninvolved conversation observation technique, recording technique, note-taking technique) and questionnaire method. The data were analyzed using the content analysis method to determine the form and percentage of language variety used in speech events on social media. In addition, Miles and Huberman's model was also used for data reduction, verification, conclusion drawing. The research found that language varieties in speech events on social media was not dominated by non-standard varieties, and inter-language mixture (code switching and code mixing). The characteristics of speech events on social media use, among others, informal language, daily-conversation language, unnatural context, and shortened writing. The results implied that the variety of language use on social media has increased since the COVID-19 pandemic. Therefore, it is necessary to research the use of codes, politeness, or hate speech on social media. Moreover, opportunities are also open to research the use of social media for language learning.

Keywords: language variety, social media, sociolinguistics, speech event

INTRODUCTION

In a multilingual and multicultural country like Indonesia, having adequate awareness of language variations across communities plays a pivotal role to maintain national identity (Martono et al., 2021). Awareness of language variations across regions is of great importance as well as an effort to establish harmony between the speakers (Putri & Sahid, 2020). This is because language awareness reflects one's sensitivity to language use and its role in human life (Finkbeiner & White, 2017; James, 1999; The Association for Language Awareness, 2012;). As commonly known, language is the community's product (Mardikantoro, 2012), which means that language or languages of a speaking

* Corresponding Author
Email haribaktim@mail.unnes.ac.id

community in its language repertoire always has variety. This implies that language can serve as not only a tool for conveying messages to other people but also a representation of the identity variation attached to the people (Sumarsono & Partana, 2002). This reinforces the importance of researching language variations in social context toward increased language awareness.

From a sociolinguistic perspective, language variation can be understood as any varieties of linguistic elements (ranging from sound to grammatical level) influenced by various factors such as different language system, identities and contexts (Hudson, 1996; Johnson & White, 2019; Wijana, 1997). As one of the main concerns in sociolinguistic studies, the term language variety (or only variety) refers to the regional, social, or contextual difference in the way certain language is used (Heydari & Barzan, 2019). Inter-language variety, dialect, and speaker are known as inter-speaker variety. Variety in the one-speaker language is called intra-speaker variety. One of the major foci of studies on language variety is to capture the picture of the non-uniformity of language users in interaction. The varieties increase when the languages used by the speaker are considerable. Therefore, variety surely accompanies each process of language change (Matsumoto, 2019); within which any language aspects (including phoneme, morpheme, syntactic structure, and meaning) can change. Hence, studies on language from sociolinguistic perspective in the present study will be approached not as language in the theoretical linguistic study but in the interaction among the community.

A portrait of studies on language variety in the Indonesian context

Indonesia, as a diverse country in linguistic backgrounds, makes its language condition unique since Indonesia is composed of many tribes and islands. This affects the language used in communication with others of the same and different ethnic backgrounds. The Indonesian language used by speakers of different ethnic backgrounds and local languages is sometimes different, and the usually used local language influences it.Several studies (Antoni et al., 2019; Iskandar, 2018; Menggo & Suastra, 2020; Qodriani & Wijana, 2020; Ramendra, 2013) have investigated the phenomenon of language variety in the Indonesian context, focusing on local languages. The major findings reveal the varieties of language use in different levels. Additionally, the studies lend support the interplay between social class and context of formality and particular use of language variant. Ramendra (2013) studied the language variety used in the speaking community of Singaraja city. The results showed that there were language varieties in the speaking community of Singaraja city related to identity of the subject. Each of the varieties tended to attach to a particular group of people, as viewed by their occupation and age.

Another similar study was conducted by Iskandar (2018) to obtain the profile of Acehnese variations from the sociolinguistic lens. The research results showed six levels of language variation, as viewed by the speakers. The six varieties identified were idiolect, dialect, sociolect, acrolect, basilect, and slang.. As viewed from the language use perspective, there were three varieties identified, namely literary variety, journalistic (non-productive) variety, and scientific variety (non-productive). From a language formality perspective, there were five varieties, namely frozen variety (non-productive), official variety (non-productive), consultative variety, causal variety, and familiar variety. From the means of communication perspective, there were two varieties: oral and written. The results implied that the Acehnese language has started to be abandoned by its speakers.

Within the Sumatran Island as in the previous study, Antoni, et al. (2019) explored language variation in Minang colloquial language among millennial citizens. Their study demonstrated that the languages used in Minang and Sijunjung had some salient differences from a phonological perspective. The changes occurred from alveolar /r/ to voiced velar fricative /gh/, from /r/ to /w/, from /a/ to /o/, and from /a/ to /aw/. The findings confirm that the language used in Minang is a standard language with phonological variety in Sijunjung. Some words may have new meanings, but many are no longer used.

Apart from the local languages, a study focusing on English language variation among Indonesian students was ever conducted by Qodriani and Wijana (2020) during the COVID-19 pandemic situation. Within the context of computer-mediated communication, language varieties were examined in 24 online classroom activities among Indonesian university students. Through textual analysis of the conversation and follow-up interviews, the results showed that the variation of the language used as a means of minimizing the disparity between teacher and learner and to maintain social relationship in digital communication.

Social media and language variety

The previous literature mentioned above reinforces that language variation is considerably influenced by the context of the communication or the situation. Hence, with the current advancement of technology for communication purposes, studies on language variation in social media interactions become more crucial than ever such as personal email communication, blog posts, and public discussions (Bamman et al., 2012; Mohammad & Yang, 2011; Volkova et al., 2013). Language variety

also appears in language use in communication interaction on social media, including WhatsApp, Instagram, Facebook, Youtube, and other types of social media. The survey was carried out in partnership with the World Economic Forum (WEF) via an online platform in June 2020 (Sebayang, 2020) The number of social media users drastically increases in line with the occurrence of the COVID-19 pandemic. Based on the survey conducted by Sathirathai on almost 70,000 youths ranging from 16-35 years old participating in the survey, 87% said they increased the use of at least one digital device during the COVID-19 pandemic.

The problem of language variation in speech events on social media is an interesting phenomenon to examine from a sociolinguistic perspective since social media exposes more dynamic and varied language than the standard language use (Lucy & Bamman, 2021). This is in line with the notion proposed by Mardikantoro (2012) that language is a product of society. The statement means that the language or languages possessed by a speech society in its language treasures always have variations. The existence of the phenomenon of using language variations in speech society is controlled by social, cultural, and situational factors (Fasold 1984, Gregory & Carroll, 2019; Hudley et al., 2020; Hudson 1996, Kartomihardjo 1981, Siregar, 2021; Wijana 1997). Social media users come from various circles with different social and cultural statuses which will bring up diverse language variations. Moreover, language variations in speech events on social media certainly have characteristics that are distinctive and different from speech events in general.

Several studies (Gouws et al., 2011; Ojetunde, 2019) have explored the language variation in social media. Gouws et al. (2011) conducted a study on contextual bearing of linguistic variation in social media used by some groups of social media users in expressing ideas through posts. The social media often used were SMS, Twitter, and Facebook. The study also revealed that different populations of users exhibited a different number of shortened forms, different English varieties, and shortening style. The results provide valuable insights into how technology can be effectively applied into the micro text. Ojetunde (2019) also conducted a similar research focusing on the varying use of deviation of English in social media interactions. The study particularly aimed to examine the extent social media language conforming standard English use. By adopting the theory of homophony and selecting messages and thirty posts on Facebook, Whatsapp, and SMS, the researcher found that language use on social media that did not conform to English use at the grammatical level, namely sentences, phrases, and words. Orthographically, some misspelled words might occur to accelerate the communication rate. Lexically, there was a literal use of new words

made with new meanings. This study concludes that there is a deviation of rules or conventions of words and sentence patterns in English that are not complied with in social media interactions.

In the Indonesian language context, several recent studies (Afdhaliyah & Haq, 2021; Hudaa & Bahtiar, 2020; Nasution et al., 2022) have investigated language variation in social media. Afdhaliyah and Haq (2021) discussed a research project exploring language variations observed on Twitter social media, specifically focusing on two languages: slang and Javanese. Using a qualitative approach through non-interactive data collection from human subjects as key informants, this study reveals language variations composed of Javanese and Indonesian language. Another study from Hudaa and Bahtiar (2020) explored how modern advancements have led to the rapid development of new vocabularies in the Indonesian language among millennial generation. The study highlights both positive and negative aspects of the emerging vocabularies, including incorrectly formed acronyms and palindromes. Fairly similar to the previous research, Nasution et al. (2022) focused on a systematic literature review examining language variations among teenagers on social media. The study identified six patterns of language variation formation, such as abbreviations, acronyms, word shortening, neologisms, and words with shifting meanings. Additionally, it categorized slang used by teenagers into positive and negative connotations. The research emphasizes the importance of understanding these language variations for effective communication on social media.

Language variety in teaching practice

The phenomenon of language variety can also be manifested in teaching practice since the students and teachers have diverse background. Several studies (Miciak et al., 2016; O Murchadha & Flynn, 2018; Stollhans, 2020; Uresin & Karakas, 2019). In the teaching of English language, Miciak et al. (2016) argue that language variety can be associated with the geographical regions or social groups or races, which influence the emergence of differing linguistic patterns among them. Since there are quite diverse social groups such as African-American, Hispanic, and Latin people, to mention a few, English at schools has varieties such as African-American English, Hispanic English, or Latino English. In a more comprehensive way, Stollhans (2020) explains that students in German often encounter language varieties and such circumstances can be beneficial for raising students' awareness and sensitivity to those varieties. Some of the factors that influence the emergence of language varieties in language learning include region, social context, degree of formality, and relationship between the speakers. Hence, students' and teachers' attitudes and orientation toward language varieties or non-

standard language has become of great attention. Murchadha and Flynn (2018), for example, posit that teachers need to have basic knowledge of certain emerging varieties in Irish context. Meanwhile, they have positive attitude and orientation toward the standard language. A similar finding was found in Uresin and Karakas' (2019) study where EFL teachers still have a strong standard language ideology in formal settings. It implicates that they have lack of awareness of the non-standard English and Turkish varieties and dialects, which might be encountered during their teaching practices and real-life language use. All in all, language variety is of great importance to be addressed not only in social context but also in instructional settings.

Although language variation in social media interactions has been quite extensively investigated, much still needs particular attention on the occurrence of language variation among Indonesian speakers in social media across situations. Since language use is related to various factors, both language and non-language factors, the present study aims to continue the discussion by exploring the characteristics of Indonesian language variation occurred in social media interactions. This research illustrates language variety in speech events on social media. There are two objectives to be revealed in this research, namely the tendency of language variety used and characteristics of language variety in speech events on social media. The research's finding is expected to be a significant new approach to developing terms of reference in fostering language awareness through examining the pedagogical implications arisen from the findings.

METHOD
Research design
This research would answer the problems of language variety use in speech events on social media, especially WhatsApp, Facebook, and Instagram. The research problems were reviewed using a qualitative approach. Specifically, the qualitative approach involved the content analysis, focused on written materials obtained from the social media. The results would be in the forms of percentages to describe and explain the trends. The qualitative approach used focused on sociolinguistic

theories. The sociolinguistic approach was used since language variety study is not only related to language structure, but also the social-cultural factors of the concerned community. With such an approach, the language variety used in social media interactions is expected to be revealed.

### Research data and data collection procedure
The data of communication snapshots in social media were obtained from the participants' responses of the questionnaire, developed by the researchers, sent via Google Forms and from the speech events in the social media used. The questionnaire has been validated by experts in this field before it was distributed to the participants. The first data sets were taken from 24 July to 4 August 2021 involving 371 participants . The participants were sorted out based on their familiarity with social media, the frequency of using social media, and the purposes of using social media. The questionnaire was used to screen the data and obtain the percentage of language variety used in the speech events on social media. The questionnaire results were supported by the speech events as the second data sets were taken from social media such as WhatsApp, Facebook, and Instagram. The research data were then screened through observing each language use. The observation was, specifically, performed by applying tapping as the basic technique, followed by advanced technique in the form of uninvolved conversation observation technique, recording technique, and note-taking technique (Sudaryanto, 2015).

Data analysis techniques
The research employed the content analysis method to examine the language variety utilized in social media interactions and determine its respective proportions. Additionally, data analysis was conducted through two main procedures: (1) analysis during data collection and (2) analysis after data collection (Miles & Huberman, 1994). The first procedure was carried out in the following steps (a) data reduction, which was to identify the language variety; (b) data presentation with matrix; and (c) making tentative conclusion/verification, either using triangulation techniques.

Figure 1
The Steps of the Second Procedure
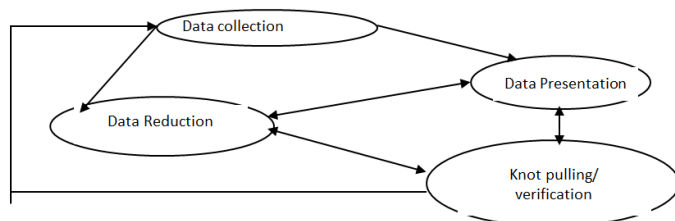
The second procedure was carried out with four main stages. First, the data collected were reduced to obtain the expected data, such as speech events on social media, particularly focused on dialect variations. For example, the datum was *betah neng kamar bae akeh sajen ne*. The datum was the result of the reduction of some Banyumas dialect data. Second, the data were presented using a matrix. Some language variations including dialects were presented as follows.

A = *Golet ndog pitik banyak/soang mbok ana sing due lur* [Look for chicken or goose eggs, maybe someone has them, brother]

B = *Betah neng kamar bae akeh sajen ne* [Just comfortable in the room, lots of food]

C = *Nek ditakoni mbojone garep kpriwe tak jawab ora ngerti nyong due duit saketemune bae ktrima alhamdullialah ora ya rpp* [If my husband asks me, I will answer that I don't know, I only have what money I have, that's acceptable thank God, no, it's okay]

At this stage, grouping or classification of data was done, based on the linguistic levels and socio-cultural factors: the datum *golet ndog pitik banyak/soang mbok ana sing due lur, betah neng kamar bae akeh sajen ne, nek ditakoni mbojone garep kpriwe tak jawab ora ngerti nyong due duit saketemune bae ktrima alhamdullialah ora ya rpp* can be grouped in dialectal language variations (Banyumas dialect). Third, the researchers interpreted the language variations, based on the language and socio-cultural factors such as dialect (Banyumas Dialect) because the word *bae* exists. Fourth, the conclusion or generalization concerned with language variation was done to accomplish the research objectives that have been formulated before.

FINDINGS AND DISCUSSION
The tendency of Language Variety Use in Speech Events on social media

Basically, in communication, humans utilize more than one language (bilingual). Thus, in speech events, the speaker can choose to use a certain language. This also occurs in speech events on social media.

Table 1 shows that in speech events on social media the participants use more than one language. In its use in a speech event, however, speakers can use one language only, two languages, three languages, or more than three languages at the same time. Such a use of language in a speech event occurs commonly. This is influenced by various factors, such as topic, situation, and hearer. Table 1 indicates that in speech events on social media,

16.5% of participants use one language, 56.2% of participants use two languages, 23.5% of participants use three languages, and 3.8% of participants use more than three languages.

Table 1
Languages Used by Participants

Number of Languages Used	Percentage
One language	16.5%
Two languages	**56.2%**
Three languages	23.5%
More than three languages	3.8%

Table 2 shows research on language used in speech events on social media. The use of the Indonesian language as the national language shows the biggest percentage of 45.6%, followed by Indonesian and local languages (code-switching or code-mixing) at 42%, Indonesia and foreign languages at the same time (code-switching or code-mixing) at 8.9% and local language of 3.5%.

Table 2
Languages Used on social media

Language Name	Percentage
Indonesian Language	45.6%
Local Language	3.5%
Indonesian and Local Languages	42%
Indonesian and Foreign Languages	8.9%

The condition shows that in speech events, speakers still prefer using the Indonesian language, even with non-standard Indonesian language variety. This can be understood since the speech event occurs in an informal situation through social media. In addition, bilingual or multilingual speakers, certainly cannot avoid using two or more languages they utilize. The use of more than one language at the same time in a speech event will lead to code-switching or code-mixing events, either with the local language or a foreign language. The speech event occurs so that what is conveyed is easier for the hearer to understand. There are certainly various factors causing code switching and code mixing to be used in a speech event. In addition, the local language is still continuously used even with the lowest percentage. This is in line with the findings of Sari et al. (2018) who state that Indonesia is a good example of the many complex relationships that can occur in the use of language, between language and identity. Language has a complex relationship with identity in Indonesia, so the regional language is still often used even though it is often mixed with Indonesian. In more detail, language variety used in speech events on social media can be observed on Table 3.

Table 3
Language Variety Use on social media

Language Variety on social media	Percentage
Non-standard Indonesian Language Variety	48.2%
Standard Indonesian Language Variety	10.6%
Slang Indonesian Language Variety	2.4%
Local Language Ngoko Variety	3.5%

Local Language Krama Variety	2.7%
Indonesia Language and Local or Foreign Language	32.5%
Foreign Language	0.1%

The tendency of language variety use in speech events on social media is presented in table 3. Table 3 shows the language variety used in speech events on social media. Indonesian language (non-standard variety) is still dominantly used by 48.2% of speakers on social media, followed by the use of Indonesian language at the same time with local language and foreign language (code switching and code mixing events) for 32.5%, Indonesian language use (standard variety) for 10.6%, local language use (ngoko variety) 3.5%, language use (krama variety) for 2.7%, Indonesian language use (slang variety) for 2.4%, and foreign language used for only 0.1%.

The results lend support the previous studies (Afdhaliyah & Haq, 2021; Hudaa & Bahtiar, 2020; Nasution et al., 2022) that social media can influence the emergence of language variations. One of the main factors of language variations emerged in the social media is the users status itself. Social media users are mostly teenagers and millennial generations where more linguistic exposure is rapid enough, proven by the use of new vocabularies in certain contexts. Such exposure becomes a consensus where understanding and using language variations is intended to achieve effective communication.

Characteristics of Language Variety in Speech Event on social media
The emergence of language variety in language practice is caused by the speakers' varied backgrounds. There is no homogeneity in the speaking community that they are composed of various backgrounds, such as differences in sex, age, education, occupation, ethnicity, and others. The dynamic situation of language use due to such diversity may lead to language variety (Lucy & Bamman, 2021; Mardikantoro, 2012).

This also occurs in speech events on social media. Speech event on social media is also performed by speaker and hearer from varied backgrounds, leading to language variety. Language variety in speech events on social media has typical and different characteristics compared to face-to-face speech events. Some characteristics of speech events in social communication include informal language use, daily conversation language use, unnatural context, and shortened writing.

Informal Language Use
Informal language is a language that is not compliant with language use principles. In written form, the language is not written in compliance with the determined principles, and in oral form, the language is daily conversation language. Informal language tends to be used in unofficial situations. Language use on social media is also categorized as language use in an unofficial situation, thus the language used is not official or informal language. Some data can be observed below:

(1) Context: Conversation in a WA group of alumni of a Senior High School (SMA) from different years.

  P1 : *Astagfirullah muka saya kenapa kaya gitu ya bundah* [Have mercy God, why does my face look like that, momma?]

  P2 : *Saya juga bingung sama muka sendiri* [I'm also confused with my face]

  P1 : *Maap bund pengen saya sensor muka saya sendiri* [Sorry momma, I want to censor my own face]

  P2 : *Ni couple goals* [This is couple goals]

  P1 : *WKWWKK gak ada obat pokoknya yg ini* [*laughing sounds* This one has no cure to it]

  P3 : *kok ya betah ya dewe ngancani duo bucin kuwi wkwkwk* [How come it's comfortable to accompany those two bucins, wk wk wk]

  P4 : *Apa lagi saya bund wkwk*

  P4 : *Ko item bgt ya dulu kita haha* [Why were we so tanned back then haha]

(2) Context: Conversation between teacher and students in a WhatsApp group.

  P1 : *Halooo selamat malam anak anak kesayanganku yang baru! Mulai hari ini aku jadi ibu tiri kalian ya selama di kelas XI ini, dulu cukup jadi guru kimia aja, sekarang jadi keluarga baru ya kita pokoe enjoy aja ya selama idup sama aku di kelas XI jan spaneng brader and sister.* [Halooo good evening my newest dearest students! Starting today I will be your stepmother throughout your 11th grade class, I use to be just your chemistry teacher, now we are a whole new family, Anyway, just be happy while living with me in class XI is very tense, bro and sister]

  P2 : *Malem bu* [Good evening, Ma'am]

  P3 : *Malam bu* [Good evening, Ma'am]

  P4 : *Malam bu* [Good evening, Ma'am]

  P5 : *Malam bu* [Good evening, Ma'am]

Data (1) and (2) are speech events on social media. The speech events in data (1) and (2) use the informal Indonesian language, as identified with the use of some words with informal language characteristics. In data (1) there are the words *kaya, maap, pengen, gak, yg, betah, dhewe, ngancani, duo. Bucin, item*, and *bgt*. In data (2) there are some words as markers of the informal Indonesian language, such as *halooo, aku, pokoe, enjoy, aja, idup, sama, jan spaneng, brader,* and *sister.*

This finding confirmed the results of Ojetunde's (2019) research that the use of language on social media is often incompatible with the formal use of language. In the use of English, the issue appears at the grammatical level of sentences, phrases, and words. Orthographically, some misspelled words may be used to speed up the pace of communication. Lexically, there is a literal use of new words made with new meanings. The study

concluded that there are deviations in the rules and conventions of a word and sentence patterns in English that are not adhered to in the use of language on social media.

Unnatural Context

Context is very important in the speech event. With context, we can analyze a speech event as what actually happens. In addition, context also determines the meaning of a speech since even the same speech with a different context will have a different meaning. Therefore, the context should be natural as actually occurring in a speech event. In speech events on social media, however, the context accompanying it is often unnatural, which means that the context is made up for the speech to be attractive. The data below can be observed:

Figure 1
Data screenshot (3)

(3) Context: An individual named Murnya Mamhnya Naura writes a status on Facebook by showing photos of food available at home
Account name: murnya mamhnya naura
Status   : *betah neng kamar bae akeh sajen ne*
[just comfortable in the room, lots of food]

Figure 2
Data screenshot (4)

(4) Context: An individual/institution nameed idntimes posts a status on social media Facebook

Status   : *pemerintah merevisi aturan vaksinasi gotong royong terutama pendanaan. Jika seblumnya vaksinasi dibebankan pada pemerintah atau badan usaha kini individu /perorangan bisa mendapatkan vaksinasi gotong royong dengan menggunakan biaya sendiri* [Government revised the gotong royong vaccination policy, especially regarding funding. Prior to this, vaccinations were burdened to the government or an institution, now individuals can get the gotong royong vaccination through personal funds.]

Data (3) and (4) are speech events on social media, especially Facebook. The contexts in the two data are unnatural, which means the contexts are intentionally created by the speakers hoping to attract hearers' or netizens (internet citizens) attention.

In data (3) an individual named Murnya Mamhnya Naura writes a status on Facebook showing photos of food available at home. The status says "*betah neng kamar bae akeh sajenne*" while showing food photos. The contexts of food photos and speech in the Javanese language of the Banyumas dialect are intentionally created to have hearers or netizen attracted. It is also possible that the food in the photos is intentionally displayed since the person making the status sells it, thus he/she actually offers the food or advertises it.

Similarly, in data (4), the context is also unnatural, where an institution with an account name idntimes makes a status on social media Facebook, "*pemerintah merevisi aturan vaksinasi gotong royong terutama pendanaan. Jika seblumnya vaksinasi dibebankan pada pemerintah atau badan usaha kini individu /perorangan bisa mendapatkan vaksinasi gotong royong dengan menggunakan biaya sendiri*". The speech context in data (4) shows a drugstore photo along with a speech. The speaker intentionally combines the speech on mutual aid vaccination funding with the name of a drugstore to attract hearers' or netizens' attention. The context is indirectly, intentionally created to highlight the photo of Kimia Farma Drugstore.

The results above reinforce the role of new media to share information, interact with each other, or exchange ideas toward the possibility of bringing out varieties in using language (Bamman et al., 2012; Mohammad & Yang, 2011; Volkova et al., 2013). Since social media today are mostly multimodal, the language patterns and choices might be different from the textual mode only where the social media users attempt to use effective and efficient language for conveying meaning to their readers or audiences. Hence, adjustments in using language through social media can influence the emergence of language varieties.

### Daily Conversation Language Use

In a speech event, there is a rule saying that language use depends on the situation. If a speech

event occurs in a formal situation, the language used should also be formal or standard Indonesian language. On the contrary, however, if a speech event occurs in an informal situation, the language used is also the informal Indonesian language.

Speech events occurring on social media can be categorized as language use in an informal situation, even if it is unnecessarily the case all the time. This causes the language used on social media to tend to be informal. The data below can be observed:

Figure 3
Data screenshot (5)
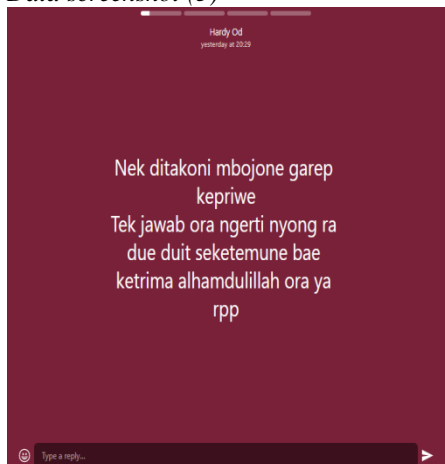
(5)　Context　:　An individual with the account name Hardy Od makes a status on social media
Status　:　*Nek ditakoni mbojone garep kepriwe tak jawab ora ngerti nyong due duit saketemune bae ktrima alhamdullialah ora ya rpp* [If my husband asks me, I will answer that I don't know, I only have what money I have, it's accepted, thank God, no, it's okay]

(6)　Context: Conversation in WA group of alumni of a Senior High School (SMA) from different years.
P1 : *Astagfirullah muka saya kenapa kaya gitu ya bundah* [Have mercy God, why does my face look like that, momma?]
P2 : *Saya juga bingung sama muka sendiri* [I'm also confused with my face]
P1 : *Maap bund pengen saya sensor muka saya sendiri* [Sorry momma, I want to censor my own face]
P2 : *Ni couple goals* [This is couple goals]
P1 : *WKWWKK gak ada obat pokoknya yg ini* [WKWWKK This one has no cure to it]
P3 : *kok ya betah ya dewe ngancani duo bucin kuwi wkwkwk* [How come it's comfortable to accompany those two bucins, wk wk wk]
P4 : *Apa lagi saya bund wkwk*
P4 : *Ko item bgt ya dulu kita haha* [Why were we so tanned back then haha]

Data (5) and (6) are speech events occurring on social media. In the speech event, the language used is daily conversation language, since the speakers

deem the speech event informal. The characteristics of informal language can be identified with the use of non-standard words, such as the Javanese language of the Banyumas dialect (data 5). In data (6), the use of informal language is on the words *kaya, gitu, sama, maap, pengen, gak, yg,* and *bgt* or the use of local (Javanese) language mixed with the Indonesian language.

Abbreviated or Shortened Word Writing
One of the symptoms of language use on social media is words abbreviation or shortening. Language users on social media who tend to be adolescents want practicality. This occurs because of adolescents' all-practical life. In language use in an informal situation, abbreviated or shortened words are indeed not prohibited, but this should be a habit since it will have an impact on language use in a formal situation, that words abbreviation and shortening is a possibility, which is to be avoided. The data to be observed:

Figure 4
Data screenshot (7)

(7)　Context　:　An indnvidual with account name @RachelVanessha twits fund on social media Twitter
Status　:　*Masi mau blg covid itu konspirasi?Masi blg covid-19 itu gaada?Masi blg yg meninggal itu cm di RS?mau aja kemakan opini hoax. sgtu bnykny yg gugur krn covid dibulan ini sampe trending innalillahi di mana2 Masi gprcy? Apa harus ngerasain kehilangan dulu baru percaya? Sakit otakmu.* [Still wanna say that covid is just a conspiracy? Still wanna say covid-19 doesn't exist? Still wanna say the ones that died are only the ones in the hospital? So easy to fell for hoaxes. So many casualties caused by covid this month to the point that innalillahi is trending everywhere. Still don't believe it? Do you have to experience losing someone first to believe it? Your brain is sick.]

Figure 5
Data screenshot (8)

(8)   Context   :   An individual with account name: @afrkml twits on social media Twiter.

Status   : *Ayo belajar lebih tegas utk nolak ajakan keluar rumah dg alasan "maaf, tdk bisa, masih pandem." Ga perlu takut! Klo kamu & keluarga tertular, belum tentu mereka mau susah2 cari IGD utkmu. Hari ini aja nimbus 20rb kasus. Jgn sampai namamu masuk di situ hny krn km ga enakan nolak.* [Let's learn to be more firm in refusing invitations to go outside by saying "Sorry, I can't, the pandemic is still happening" No need to be afraid! If you and your family are infected, there's no guarantee that they want to go to lengths to look for ERs for you. Today alone, it had reached 20K cases. Don't let your name get on that list just because you didn't had the heart to refuse.]

Data (7) and (8) are speech events on social media, especially Twiter. In the speech events, there is much of language with intentional abbreviation or shortening.

Data (7) are of language use on social media Twiter with abbreviation and shortening. An individual with account name @RachelVanessha twits on social media Twitter, saying "*Masi mau blg covid itu konspirasi? Masi blg covid-19 itu gaada? Masi blg yg meninggal itu cm di RS? mau aja kemakan opini hoax. Sgtu bnykny yg gugur krn covid dibulan ini sampe trending innalillahi di mana2 Masi gprcy? Apa harus ngerasain kehilangan dulu baru percaya? Sakit otakmu.*" [Still wanna say that covid is just a conspiracy? Still wanna say covid-19 doesn't exist? Still wanna say the ones that died are only the ones in the hospital? So easy to fell for hoaxes. So many casualties caused by covid this month to the point that innalillahi is trending everywhere. Still don't believe it? Do you have to experience losing someone first to believe it? Your brain is sick.] In the data (7), there are some words that are abbreviated or shortened. The abbreviated words include *blg* (bilang), *cm* (cuma), *sgtu* (segitu), *bnykny* (banyaknya), *yg* (yang), and *gprcy* (gak percaya).

The shortened words include *masi* (masih) and *gaada* (tidak ada).

Data (8) are of language use on social media, to be exact Twiter. An individual with account name: @afrkml twits on social media Twiter saying "*Ayo belajar lebih tegas utk nolak ajakan keluar rumah dg alasan "maaf, tdk bisa, masih pandemi. Ga perlu takut! Klo kamu & keluarga tertular, belum tentu mereka mau susah2 cari IGD utkmu. Hari ini aja nembus 20 rb kasus. Jgn sampai namamu masuk di situ hny krn km ga enakan nolak*" [Let's learn to be more firm in refusing invitations to go outside by saying "Sorry, I can't, the pandemic is still happening" No need to be afraid! If you and your family are infected, there's no guarantee that they want to go to lengths to look for ERs for you. Today alone, it had reached 20K cases. Don't let your name get on that list just because you didn't had the heart to refuse]. In data (8) there are abbreviated and shortened words. The abbreviated words include *utk (untuk), dg (dengan), tdk (tidak), rb (ribu), jgn (jangan), hny (hanya), krn (karena),* and *km (kamu)*. The shortened words include *ga (gak, tidak),* and *aja (saja).*

Implications to language teaching
The language variety occurred in the social interactions identified in the present study raises issues on how such cultural contents can be meaningful learning resources in the Indonesian language classrooms. First, teachers need to have adequate awareness of how students' race can influence their strategies to teach Indonesian language. A review made by Von Esch et al. (2020) pinpoints the role of race in shaping standards and hierarchies of languages. It means that the teachers must be concerned with various students' races in transmitting the linguistic contents of Indonesian language in the classroom, for instance, being aware of students' dialects in pronouncing Indonesian words, particularly in the suburban areas. Second, teacher identity cannot be neglected in viewing the teaching and learning of language variety. Teachers' ethnicity, gender, culture, and language may have impact on how they address and deal with students' interests, needs, and concerns in learning the language (Taylor, 2017). For example, the linguistic history and background of the teacher may cause unequal attention to language varieties outside their regions, so that students whose linguistic background is different from their teachers may feel excluded and not prioritized to be concerned with. Third, in a multilingual setting like Indonesia, it is pivotal to incorporate communicative competence in assessing students' language skills. Hence, teachers as classroom designers are encouraged to formulate a comprehensive learning process.

CONCLUSION

This study has investigated the characteristics of Indonesian language variety in speech events occurred in social media. The results generally imply that language variety is the picture of non-uniformity of language users in interaction. Language variety also emerges in language user in speech event on social media, including WhatsApp, Instagram, and Facebook Language variety in speech event on social media tends to be dominated by informal variety, code switching and code mixing). The characteristics in speech event on social media include use of informal language, use of daily life conversation language, unnatural context, and shortened writing. The results of the present study can be used as a reference for further examining the nature of Indonesian language variation, which particularly has implication to classroom discourse. Since students are also social media users, their language use in the classroom context might be influenced by the language used in the social media. Hence, teachers need to raise language awareness of the students to better prepare them to be fluent and good Indonesian language users. Therefore, it is necessary for further research to explore the use of codes, politeness, or hate speech on social media. Moreover, opportunities are also open to research the use of social media for language learning.

REFERENCES

Afdhaliyah, R., & Haq, S. C. (2021). Variasi bahasa dalam media sosial Twitter. *GHÂNCARAN: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, *3*(1), 104-116.

Antoni, C., Irham, I., & Ronsi, G. (2019). Language variation in Minang colloquial language spoken in Kabun Region: Sociolinguistic study on millennial citizens. *Jurnal Arbitrer*, *6*(2), 92-98.

Association for Language Awareness (ALA). (2012). About language awareness. Association for language awareness. http://www. languageawareness.org/web.ala/web/about/tout .php. Accessed 6 Feb 2015.

Bamman, D., Eisenstein, J., & Schnoebelen, T. (2012). Gender in Twitter: Styles, stances, and social networks. https://arxiv.org/vc/arxiv/papers/1210/1210.45 67v1.pdf

Fasold, R. (1984). *The sociolinguistics of society*. Basil Blackwell Publisher.

Finkbeiner, C., & White, J. (2017). Language awareness and multilingualism: A historical overview. In J. Cenoz, D. Gorter & S. May (Eds.), *Language Awareness and Multilingualism Encyclopedia of Language and Education* (pp. 3-17). Springer International Publishing.

Gouws, S., Metzler, D., Cai, C., & Hovy, E. (2011, June). Contextual bearing on linguistic variation in social media. In *Proceedings of the workshop on language in social media (LSM 2011)* (pp. 20-29). Portland, Oregon: Association for Computational Linguistics

Gregory, M., & Carroll, S. (2018). *Language and situation: Language varieties and their social contexts*. Routledge.

Heydari, B., & Barzan, P. (2019). *Language variation*. https://www.researchgate.net/publication/3374 99049

Hudaa, S., & Bahtiar, A. (2020). Variasi bahasa kaum milenial: Bentuk akronim dan palindrom dalam media sosial. *ESTETIK: Jurnal Bahasa Indonesia, 3*(1), 41-52.

Hudley, A. H. C., Mallinson, C., & Bucholtz, M. (2020). Toward racial justice in linguistics: Interdisciplinary insights into theorizing race in the discipline and diversifying the profession. *Language*, *96*(4), e200-e235.

Hudson, R. A. (1996). *Sociolinguistics (2nd Ed.)*. Cambridge University Press.

Iskandar, D., Pujiono, M., & Samad, I. A. (2018). The profile of Acehnese variation: Sociolinguistic analysis. *International Journal of Comparative Literature and Translation Studies*, *6*(3), 12-19.

James, C. (1999). Language awareness: Implications for the language curriculum. *Language, Culture and Curriculum, 12*(1), 94–115 https://doi.org/10.1080/07908319908666571

Johnson, E. K., & White, K. S. (2019). Developmental sociolinguistics: Children's acquisition of language variation. *WIREs Cognitive Science, 11*(1), e1515. https://doi.org/10.1002/wcs.1515

Kartomihardjo, S. (1981). *Ethnography of communicative codes in East Java*. Dept. of Linguistics, Research School of Pacific Studies, The Australian National University.

Lucy, L., & Bamman, D. (2021). Characterizing English variation across social media communities with BERT. *Transactions of the Association for Computational Linguistics*, *9*, 538-556. https://doi.org/10.1162/tacl_a_00383

Mardikantoro, Hari Bakti. (2012). Pilihan bahasa masyarakat Samin dalam ranah keluarga (Language Choices of Samin Communities in Family Domains). *Jurnal Humaniora, 24*(3), 345-357.

Martono, M., Dewantara, J. A., Efriani, E., & Prasetiyo, W. H. (2022). The national identity on the border: Indonesian language awareness and attitudes through multi-ethnic community involvement. *Journal of Community*

Psychology, *50*(1), 111-125. https://doi.org/10.1002/jcop.22505

Matsumoto, K. (2019). Language variation and change. In P. Heinrich & Y. Ohara (Eds.) *Routledge handbook of Japanese Sociolinguistics* (pp. 199-217). Routledge.

Menggo, S., & Suastra, I. M. (2020). Language use and language attitudes of Sumbawanese speakers in Bali. *Register Journal, 13*(2), 333-350. https:// doi.org/10.18326/rgt.v13i2.333-350

Miciak, J., Wilkinson, C., Alexander, C., & Reyes, P. (2016). Addressing language variety in educational settings: Toward a policy and research agenda. *Educational Policy*, *30*(6), 820-848.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage Publications.

Mohammad, S. M., & Yang, T. W. (2011). Tracking sentiment in mail: How genders differ on emotional axes. *In Proceedings of the ACL 2011 Workshop on Computational Approaches to Subjectivity and Sentiment Analysis WASSA, Portland, OR*.

Nasution, F., Siregar, A. K., Sopiyanto, S. F., & Lubis, N. S. (2022). Systematic literature review: Variasi bahasa pada remaja dalam penggunaan media sosial. *EduPsyCouns Journal: Journal of Education, Psychology, and Counseling, 4*(2), 154-161. https://doi.org/10.33487/edupsycouns.v4

Ojetunde, C. F. (2019). The use of English on social media: Deviation of variation. The European Conference on Arts & Humanities

Ó Murchadha, N. O. E. L., & Flynn, C. J. (2018). Educators' target language varieties for language learners: Orientation toward 'native' and 'nonnative' norms in a minority language context. *The Modern Language Journal*, *102*(4), 797-813.

Putri, A. D. F. E., & Sahid, M. (2020). Building regional language awareness through personal and social communication as an acculturation efforts. *Palakka: Media and Islamic Communication*, *1*(1), 90-99.

Qodriani, L. U., & Wijana, I. D. P. (2020). "Drop your 'Hello!' here!": Investigating the language variation used in online classroom for tertiary level in Indonesia. *Proceedings of the International Joint Conference on Arts and Humanities (IJCAH 2020), 491,* 617-623.

Ramendra, D. P. (2013). Variasi pemakaian bahasa pada masyarakat tutur kota Singaraja. *Jurnal Ilmu Sosial dan Humaniora*, *2*(2), 275 - 287.

Sari, B. T., Chasiotis, A., van de Vijver, F. J., & Bender, M. (2018). Parental culture maintenance, bilingualism, identity, and well-being in Javanese, Batak, and Chinese adolescents in Indonesia. *Journal of Multilingual and Multicultural Development*, *39*(10), 853-867.

Sebayang, R. (2020). *Di tengah pandemi, penggunaan teknologi meningkat pesat* [In the middle of pandemic, the utilization of technology has rapidly developed]. CNBC Indonesia. https://www.cnbcindonesia.com/tech/2020072 3190647-37-175009/di-tengah-pandemi-penggunaan-teknologi-meningkat-pesat

Siregar, I. (2021). Analysis of Betawi Language Interference on the Morphology of Adolescent Speech in Jakarta. *Journal of Humanities and Social Sciences Studies*, *3*(8), 54-60. https://doi.org/10.32996/jhsss.2021.3.8.7

Stollhans, S. (2020). Linguistic variation in language learning classrooms: considering the role of regional variation and 'non-standard' varieties. Retrieved from http://www.meits.org/policy-papers/paper/linguistic-variation-in-language-learning-classrooms-considering-the-role-of-regional-variation-and-non-standard-varieties

Sudaryanto. (2015). *Metode dan aneka teknik analisis bahasa [Methods and techniques in analyzing language]*. Sanata Dharma University Press.

Sumarsono & Partana, P. (2002). *Sosiolinguistik (Sociolinguistics)*. Sabda.

Taylor, A. (2017). Putting race on the table: How teachers make sense of the role of race in their practice. *Harvard Educational Review, 87*(1), 50–73, 157

Üresin, F., & Karakaş, A. (2019). Investigation of Turkish EFL teachers' views about standard languages, dialects and language varieties through the lenses of English and Turkish. *The Literacy Trek*, *5*(2), 1-24.

Volkova, S., Wilson, T., Yarowsky D. (2013). Exploring demographic language variations to improve multilingual sentiment analysis in social media. *In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, (pp 1815–1827).

Von Esch, K. S., Motha, S., & Kubota, R. (2020). Race and language teaching. *Language Teaching*, *53*(4), 391-421.

Wijana, I. D. P. (1997). *Linguistik, sosiolinguistik, dan pragmatik (Linguistics, sociolinguistics, and pragmatics)*. Paper presented in Scientific Simposium of Language and Culture in Balai Bahasa Yogyakarta.