ATLANTIS
PRESS

Utilization of the Corpus in Depicting Impoliteness in Indonesian on Social Media

Miftahulkhairah Anwar[1,*] Fathiaty Murtadho[2,] Herlina Usman[3,] Reny Heryanti[4]
Fachrur Razi Amir[5]

[1,2,3] *Unversitas Negeri Jakarta*
[4] *Universitas Jambi*
[5] *Universitas Djuanda Bogor*
[*]*Corresponding author. Email:* miftahulkhairah@unj.ac.id

ABSTRACT

The use of social media is massively increased, but it has not been accompanied by the awareness of proper and polite language. As reported by the Tempo National Page, there were 476 complaints of hate speech on social media that were submitted to the Cyber Crimes of the Criminal Investigation Agency of the Police from February 23 to May 11, 2021. This hate speech is a portrait of the lack of awareness of social media users to speak politely and the lack of awareness not to neglect the rights and honours of others. Language impoliteness on social media has destructive power, not only its spread is very fast, but it even has the potential to threaten the integrity of the nation. Therefore, this paper aims to map impolite words on social media status through the use of a digital corpus. By using qualitative descriptive research methods, it was found that there are two categories of impolite forms commonly used by netizens: first, negative connotations, 153 lexemes and second, potentially impolite neutral connotations, 26 lexemes. This form mapping can be used as a trigger so that social media users are aware of sorting, choosing, and using diction correctly.

Keywords: Language Impoliteness, Corpus Linguistics, Indonesian Language, Social Media.

1. INTRODUCTION

The presence of information technology has changed the way humans communicate from the real world to the virtual one. Dissemination of information is no longer dominated by conventional media, but has been replaced by new social media. The presence of social media is becoming a trend in all circles because it offers convenience and speed in interacting. The rise of social media has an impact on the quality of information. Now, everyone can become a maker and disseminator of information without taking into account the social effects caused. Manduric [1] in his article "*Social Media as a Tool for Information Warfare*" states that social media acts as a weapon of mass destruction and triggers conflict.

Reporting from the Tempo National Page [2], there were 476 complaints of social media hate speech that were submitted to the Cyber Crimes of the Criminal Investigation Police during the period from February 23 to May 11, 2021. The social media that received the most warnings was Facebook with a total of 228 accounts. Then, Twitter with 224 accounts, Instagram with 14 accounts, YouTube with nine accounts and WhatsApp with one account. Then, there were 70 contents were removed, 179 submitted for blocking,

and 67 more still pending. According to Rangkuti and Lubis [3], problems that can threaten the integrity of the nation do not only come from the differences in political ideological attitudes, but can also come from language impoliteness. This paper focuses on the form of language impoliteness in the Facebook status account as the social media most given warning.

There have been many studies on politeness and impoliteness in language, including Xiang, et al. [4], Li, et al. [5], Baider [6], Alahmad, et al. [7], Wong & Esler [8], Astia [9], Darong and Basthomi, [10], Fitriyani and Andriyanti [11], Widagdo, et al. [12], Humaizi, et al. [13], Mahmud [14], Palupi and Endahwati [15], Foolen [16], Nugrahani [17], Culpeper, Haugh, and Kadar [18], Prakash and Kumar [19], Rasyikin [20], Olorunleke, Obidiran , and Mustafa [21], Tretyakova [22], Anwar [23] Anwar [24], Anwar [25], Anwar [26].

Language politeness as part of a pragmatic study discusses the relationship between language and something outside the language. This paper is relevant to the research conducted by Culpeper and Hardaker [18] which puts the analysis of language impoliteness into three concepts, namely: context, power, and impoliteness formulation. This paper is also relevant

and is a continuation of what Anwar [27] and Anwar [28] have done. Anwar [27] and Anwar [28] show that language impoliteness markers on social media are characterized by diction: 1) ignoring the interlocutor, 2) separating themselves, (3) being unsympathetic, 4) using inappropriate markers/designations, 5) using a secret language that is not understood by the interlocutor, 6) using taboo, rude, or profane language, 7) using derogatory nicknames in greeting, 8) frightening, 9) demeaning/harassing, 10) ridiculing or mocking, 11) insulting, 12) not treating partners seriously, 13) attacking others, 14) using negative pronouns, 15) placing another person in charge, 16) pretending polite, 17) not thanking partners, (18) not congratulating, 19) illogical speech, 20) making a status lie, 21) identical to the animal world, 22) synonymous with goods, 23) mixed variety of languages, 24) showed disapproval, 25) ordered the interlocutor in another way directly, 26) showing jealousy, 27) showing lust, 28) interrupting, 29) showing disgust. Therefore, as a continuation of what has been done in previous research, this paper focuses on the use of the linguistic corpus to emphasize and strengthen the sense value of potentially impolite words. The corpus used is corpora.uni-leipzig.

2. RESEARCH METHOD

This research was conducted by following the working principles of descriptive research and qualitative design. The data in this study is a linguistic unit that shows language impoliteness on *Facebook* status in the November-December 2019 period. The data collection used is the listening method with the note-taking technique, namely recording the data obtained in data cards. The series of data analysis activities are: 1) data reduction which includes identification, classification, and coding by utilizing the corpora.uni-leipzig device; 2) speech data analysis; and 3) verification of data interpretation.

3. FINDINGS AND DISCUSSION

Observations on the use of language on social media during November-December 2019 showed that there were 179 impolite lexemes commonly used by netizens. This form is divided into two categories: (1) words with negative connotations and (2) neutral words with the potential to be disrespectful. This grouping, beside based on research by Anwar [27] and Anwar [28] is also based on the word node in the corpus. Words with negative connotations were found as many as 153 words or 85.5%, while those with neutral connotations found as many as 26 words or 14.5%.

3.1 Words with Negative Connotations Impoliteness

Forms of impoliteness on Facebook are mostly characterized by the use of words with negative connotations. Here are some examples of Facebook statuses that contain words with negative connotations.

(1)   Dasar *Bajingan*! (*Bastard*!)

The word *bajingan* in example (1) has a negative nuance and contains swearing. If someone's status contains the word *bajingan*, then that status tends to be positioned impolite. This is due to the negative connotation inherent in the word *bajingan*. The search for the word collocation *bajingan* on corpora.uni-leipzig also shows negative nuances, as shown in Figure 1 below.

Figure 1 The results of the search for the words *bajingan* corpora.uni-leipzig

(2)   *"Ternyata isinya orang-orang **penipu, bangsat, pembegal** uang jamaah haji."* (It turns out that the contents are *penipu*, *bangsat*, *pembegal* (fraudsters, bastards, muggers) of pilgrims' money.

This status contains the words *penipu, bangsat,* and *pembegal* so that when used in social media status, the status is positioned as a form of disrespect because it shows directly and openly to the public about negative behavior committed by someone. This has the potential to threaten the face and demean a person's dignity. The inherent meaning of these three words has a negative connotation. word *penipu* For example, the searching for the collocation of the word *penipu* on corpora.uni-leipzig shows negative nuances, as shown in the following image.

Figure 2. Search Results for **penipu** on corpora.uni-leipzig

- Namun waspadalah karena banyak **penipu** berkeliaran di media berbagi foto milik Facebook ini. (www.harianhaluan.com, collected on 15/07/2020)
- Dua orang **penipu** dari negara bagian Uttar Pradesh di India ditangkap pada pekan ini. (lifestyle.okezone.com, collected on 02/11/2020)
- Namun, ternyata itu hanya akal-akalan si **penipu** hingga membuat Baim geram. (www.tribunnews.com, collected on 06/03/2020)
- Begini akal bulus komplotan **penipu** yang diringkus di Tol Sragen untuk kabur dari petugas. (www.solopos.com, collected on 13/07/2020)
- Tak hanya itu, Kaesang juga sempat menjahili salah satu akun jualan online **penipu**. (www.harianhaluan.com, collected on 10/09/2020)
- Pihaknya saat ini masih menelusuri apakah ada pondok pesantren yang terpedaya dengan ulah oknum **penipu**. (www.kpogja.com, collected on 24/09/2020)
- Saat ada kesempatan, **penipu** tersebut akan menggores bodi mobil, lalu menawarkan bantuan. (www.viva.co.id, collected on 13/01/2020)

(3) …org yg blg "X tuh **licik**, perhitungan suaranya bisa dimainin"… (…. people who say "X is *cunning*, the vote count can be modified"….)

The word *licik* on the status inherently negative connotations. *Licik* (cunning) means good at deceiving, a lot of bad sense, deceit, and cheating. This is in line with the word collocation *licik* following.

Figure 3. Search ResultsPassword **licik** on corpora.uni-leipzig

- Selain sadis, pelaku juga terkenal **licik** dan lincah untuk menghindar dari kejaran polisi. (www.harianhaluan.com, collected on 07/06/2020)
- Anda bijaksana, tetapi juga bisa **licik**. (www.tribunnews.com, collected on 20/03/2020)
- Tak hanya sadis dan kejam, Totok Dwi Prasetyo juga **licik**. (www.tribunnews.com, collected on 26/02/2020)
- Anjing sedang tak kenal kompromi dan sedikit **licik** minggu ini. (www.tribunnews.com, collected on 14/04/2020)
- Kamu bijaksana, tetapi juga bisa **licik**. (www.tribunnews.com, collected on 20/03/2020)
- Aries bijaksana, tetapi juga bisa bersikap **licik**. (www.tribunnews.com, collected on 21/03/2020)
- Namun, ada solusi yang sedikit **licik** yang memungkinkan untuk melakukannya di browser desktop. (www.harianhaluan.com, collected on 29/07/2020)
- Pelaku yang tidak tamat sekolah dasar ini punya modus **licik**. (www.jpnn.com, collected on 12/06/2020)
- Laku, biasanya besarnya bunga yang sangat tinggi bisa ditutup-tutupi oleh penyalur mobil yang **licik**. (economy.okezone.com, collected on 09/02/2020)
- Bahkan keluarga ini sampai dituding sebagai tukang tipu dan disebut penipu **licik**. (bangka.tribunnews.com, collected on 30/08/2020)

(4) *Udah gk tau malu... mau menang dengan cara2 biadab !!!!!*

The word **biadab** on the status is inherently negative connotations. **Biadab** means ignorant, insolent, and cruel. This is in line with the following *biadab* word collocation.

Figure 4. Search Result for *Biada*b on corpora.uni-leipzig.

- Ia menyebut Budi, **biadab**. (kupang.tribunnews.com, collected on 27/02/2020)
- Ai Maryati menyebut kasus tersebut sebagai serangkaian kekerasan yang **biadab** dan merusak generasi muda. (www.tribunnews.com, collected on 21/01/2020)
- Menurut Cucun, tindakan Ali Kalora membunuh empat orang yang masih satu keluarga jelas salah dan **biadab**. (www.suarasurabaya.net, collected on 30/11/2020)
- Nahasnya, sang pacar tak melakukan tindakan **biadab** itu seorang diri. (pontianak.tribunnews.com, collected on 27/03/2020)
- Kedelapan pemuda **biadab** itu sudah ditangkap polisi dan ditahan. (www.viva.co.id, collected on 08/07/2020)
- Aksi **biadab** yang dilakukan Suherman masih membekas di ingatan para wanita yang menjadi korbannya. (www.tribunnews.com, collected on 05/08/2020)
- Mantan istri Zack Lee ini mengaku sakit hati pada para pelaku yang menurutnya **biadab**. (celebrity.okezone.com, collected on 16/04/2020)
- Ia dengan keras mengutuk tindakan keji dan **biadab** tersebut, serta telah menyampaikan belasungkawa kepada keluarga Cisse. (www.viva.co.id, collected on 21/08/2020)

Apart from these four examples, negative connotations were also found which are commonly used by netizens on Facebook. The words are *banci, sinting, kapir, sotoy, bangkotan, arogan, stres, ngibul, dungu, bajingan, keparat, dungu, bego, gampar, kampret, bahlul, biadab, jancuk, tolol, genderuwo, sontoloyo, goblok, dungu, iblis, bangsat, bacot, munafik, pelakor, najis, jablay, laknat, penipu, pembegal, setan, bodoh, sok, ngeles, nyinyir, anjrit, sableng, abal-abal, brengsek, bajingan, idiot, planga-plongo, kacung, lonte, pecundang, pembohong, culas, cemen, bego, sialan, bloon, gila, jelek, jongos, penjilat, kunyuk, biadab, sombong, busuk, sinting, perek, oon, mampus, laknat, jijik, bedebah, sedeng, pengecut, pelacur, bencong, titit, pengkhianat, memalukan, pekok,* selangkangan, *geblek, modar, brutal, kebobrokan, kedok, jahanam, buset,*

licik, mesum, murahan, koplak, tabok, cabul, fitnah, picek, terkutuk, penista, culun, bullshit, gila, tablo, gedeg, jalang, alay, badut, bala, sampah, anjing, cebong, bantet, jeplak, tai, suek, bangkai, budeg, lemot, monyet, binatang, sempak, bala, korengan, payah, homo, benalu, tengik, gendut, bunuh, keok, kotoran, menggonggong, asu, daki, kurus, kering, jamban, haram, parno, nyet, cukong-cukong, nyungsep, jahannam, rakus, edan, daki, tonggos, juling, nyungsep, pocong, lemah, lemot, kutukan, cebol, cebokin, and *jigong.*

Inherently, these words have a negative connotation so that their appearance in Facebook status implies impoliteness in language. A search of corpora.uni-leipzig shows that even these words are mostly collocated with words with negative connotations.

3.2 Neutral Words Potentially Impoliteness

Language on Facebook is also marked by the use of words that inherently mean neutral, but have the potential to be disrespectful because they are coupled with other words in a construction. The following are some examples of statuses on Facebook timeline that contain neutral, but potentially impolite words.

(5) *Heran,* Anda masih mau mendukung calon yang seperti ini... *narasi*! (Surprised, you still want to support a candidate like this… *narrative*!)

Narasi is the telling of a story. This word does not have a negative connotation. The corpora.uni-leipzig search shows that collocatively, *narasi* are usually accompanied by neutral words. However, the presence of the word *heran* in this status gives a level of disrespect.

Figure 5. Search Results for Words *narasi* on corpora.uni-leipzig

(6) Cis, Mukanya *kentang* banget! (Cis, her face is very *potato*!)

Kentang is *a* type of plant that contains a lot of starch and is favored by almost all circles. However, the presence of the word *cis* as an interjection of disgust has the potential to eliminate the neutral value of the word *kentang* in that status. In fact, in collocation in the following corpus, *kentang* are side by side with the word neutral.

Figure 6. Search Results for words *kentang* on corpora.uni-leipzig

(7) Hai *Onta*, apakah masih bisa mendengar jerit penderitaan warganya? (Hi *Camel*, can you still hear the suffering screams of your citizens?)

The standard form of the word *onta* is *unta*, an animal that lives in Arabia. The corpus search shows that collocatively, *unta* is always side by side with religious terms. However, the presence of the word *penderitaan* (suffering) in this status reduces the neutral value of the word *unta*. In this status, *unta* becomes a negative nickname for a certain group.

Figure 7. Results of Searching for *unta* in corpora.uni-leipzig

Apart from these three examples, neutral words were also found, but potentially impolite due to the context of the sentence. Words *boneka, wawasan, rekening, tukang, muka, otak, jenderal, tua, penguasa, mulut, batang, kaleng, kambing, onta, monyet, ngotak, plastik, udang, babi, cebong, badut, kentang, anjing, cari, wanita,* and *kambing* does not inherently have a negative connotation. However, the presence of a certain word that accompanies it in a construction has the potential to eliminate the neutrality of the word.

4. CONCLUSION

The use of digital corpus in depicting language impoliteness on social media reinforces the characteristics of impolite form of politeness. Furthermore, this constancy has the potential to be used as a basis for developing a status filtering design with impolite content on social media.

AUTHORS' CONTRIBUTIONS

The title "AUTHORS' CONTRIBUTIONS" should be in all caps.

ACKNOWLEDGMENTS

REFERENCES

[1] Manduric, Aylin. (2016). *Social Media as a Tool for Information Warfare.* New York: Springer.

[2] Amirullah and Andita Rahmah. (2021). "Virtual Police Send Alerts to 336 Accounts As of May 11, Most Facebook". *National Tempo.* Published May 14, 2021.

[3] Rangkuti, Rahmadsyah and Lubis, Andi Pratama. (2018). "Problems in Multicultural Society: From Language Politness to Hate Speech". *Proceeding of the Ist Annual International Conference on Language and Literature, 18-19 April 2018*, Medan Indonesia, 255-261.

[4] Xiang, Xia, Binghan Zheng, Dezheng Feng. (2020). Interpreting impoliteness and over-politeness: An investigation into interpreters' cognitive effort, coping strategies and their effects. *Journal of Pragmatics*, Vol. 169, 231-244.

[5] Li, Mingyang, Louis Hickman, Louis Tay, Lyle Ungar, Sharath Chandra Guntuu. (2020). Studying Politeness across Cultures using English Twitter and Mandarin Weibo. *Proceedings of the ACM on Human-Computer Interaction,* Vol. 4, No. CSCW2.

[6] Baider, Fabienne H. (2020). Researching Politeness: From the 'Classical' Approach to Discourse Analysis … and Back. Corpus Pragmatics 4*: International Journal of Corpus Linguistics and Pragmatics*, 259-272.

[7] Alahmad, *Tarek Hider Mohammad and Asma Khaled Abdullah Alkasassbeh. (2020).* Linguistic Politeness and Gender: Apology Strategies: A Sociolinguistic Research. *International Journal of Linguistics*, Vo. 12, No.1, 167-178.

[8] Wong, L., & Esler, J.. (2020). A Review of Teaching and Learning Linguistic Politeness. *Journal of Critical Studies in Language and Literature*, *1*(4), 13-21.

[9] Astia, Idda. (2020). Politeness Strategy in Interlanguage Pragmatics of Complaints by International Students. *Indonesian Journal of English Language Teaching and Applied Linguistics*. Vol 4(2), 349-362.

[10] Canggu Darong, H., Kadarisman, AE., & Basthomi, Y. (2020). Teachers' Politeness Markers in Request in Classroom Interactions. *NOBEL: Journal of Literature and Language Teaching*, *11*(2), 217-233.

[11] Fitriyani, Santi & Erna Andriyanti. (2020). Teacher and Students' Politeness Strategies in EFL Classroom Interactions. *Indonesian Journal of English Language Teaching and Applied Linguistics*, Vol 4 (2), 259-273.

[12] Widagdo, Titis Bayu, et al. "Strategy Politeness of Werkudara in Wayang purwa." *International JJournal of Linguistics, Literature and* Translation, vol. 3, no. 6, 15 Jun. 2020, pp. 127-134

[13] Humaizi, Siti Hazzah Nur Ritonga, Sakhyan Asmara, Muhammad Yusuf. (2020). Positive and Negative Politeness Strategies Used in the House of Representative Members' Discourse: A Case Study in Medan, Indonesia. *International Journal of Advanced Science and Technology*, *29*(05), 6391 - 6399.

[14] Mahmud, Pure. (2019). "The Use of Politness

strategies in The Classroom Contex by English University Student". *Indonesian Journal of Applied Linguistics*, 8(3): 597-606.

[15] Palupi, Muncar Tysa and Endahwati, Nafisah. (2019). "Language Politeness in Online Social Media: A Descriptive Review on Political News Commentary on Facebook". *Journal of Scripts*, 5(1): 26-31.

[16] Foolen, Ad. (2019). "Qua Vadis Pragmatics?From Adaptation to Participatory Sense-Making". *Journal of Pragmatics*, 145:39-46.

[17] Nugrahani, Farida. 2017. "The Use of Language in Social Media and Its Implications for the Character of the Nation". *Journal of Stylistics*, Vol.3, No.1 of 2017, pp. 1-16.

[18] Culpeper, Jonathan and Hardaker, Claire, (2017). Impoliteness, in Culpeper, Haugh, Kadar (Ed.), *The Palgrave Handbook of Linguistics (Im)politeness*, (pp.215-220). United Kingdom: Macmillan Publisher.

[19] Prakash, Om and Kumar, Rajesh. (2017). "Linguistics (Im)politeness and Public Discourse in Media Sphere. *International Journal of Innovations in Tesol and Applied Linguistics"*; 3(1): 1-12.

[20] Rashikin, Chairil. (2018). "Differences from Politeness Principles in Indonesian in SMP Negeri 2 Tampiala Village, South Dampal District, Toli-Toli Regency". *Journal of Language and Literature*; 3(5):1-10.

[21] Olorunleke, Sunday Femi; Obidiran, Gbenga; and Mustafa, Linda Jummai. (2017). "Pragmatic Analysis of Invective Language on Social Media". *International Journal of Innovative Research and Development*; 6 (10): 128-134.

[22] Tatyana P. Tretyakova. (2016). *"*On Politeness in Translation*" Journal of Siberian Federal University. Humanities & Social Sciences 3 (9),* 653.

[23] Anwar, M. (2013). *Optimization of Language Engineering: Towards a Logical and Courteous Mass Media*.

[24] Anwar, M. (2014). Dignifying the Nation through the Use of Polite and Logical Mass Media Language. *International Congress of the Indonesian Linguistic Society (KIMLI)*.

[25] Anwar, M. (2020). *The Impoliteness of Indonesian Language on Instagram Potentially toward Conflict of Ethnic, Religion, Race and Inter-group*. https://doi.org/10.5220/0009940021342141

[26] Anwar, M., Murtadho, F., & Razi Amir, F. (2020). *Understanding the Poster Texts of the Student's Demonstration Activities: Perspective of Language Impoliteness*.

[27] Anwar, M. (2018). *Development of Indonesian Language Politeness Model in Social Media Based on Functional Linguistics as an Effort to Prevent SARA Conflict*.

[28] Anwar, M. (2019). *Development of Indonesian Language Politeness Model in Social Media Based on Functional Linguistics as an Effort to Prevent SARA Conflict*.