ANALYSING SOCIAL PRESENCE IN WHATSAPP APPLICATION USING LIWC ANALYSIS

Siti Norzaimalina Abd Majid, Abdul Rahim Salam
Universiti Malaysia Pahang Al-Sultan Abdullah
norzaimalina@umpsa.edu.my
Universiti Teknologi Malaysia
m-arahim@utm.my

ABSTRACT

Instant messaging platforms such as WhatsApp has become a popular mode of communication adopted in online learning. A high level of social presence is essential for online learning to be effective. While many studies have explored social presence in various online contexts, there is a gap in understanding the linguistic elements that indicate social presence using computerised text analysis programs like Linguistic Inquiry and Word Count (LIWC). Therefore, this study aims to analyze social presence, specifically in WhatsApp, using the LIWC program. Seven linguistic categories from LIWC, namely clout, authentic, tone, negative emotions, positive emotions, social processes, and first pronoun (I, We), were chosen as social presence indicators. WhatsApp conversations from four groups of postgraduate students were selected in this study. Generally, the average word count for a 17-week conversation was 3861 words per group. The groups showed a high confidence level and expertise (clout) in WhatsApp conversations. However, they used a distanced form of discourse in their messages (authentic). The conversational tone was upbeat style and positive (tone). The first pronoun 'we' was used more than "I". About 7% of overall messages were positive, and only 1% of negative emotions. Meanwhile, social processes portrayed were about 13% in general. This study highlights that linguistic elements can be used to imply the existence of social presence in WhatsApp, leading to more data-driven insights. Furthermore, educators and instructional designers can enhance the online learning experience by incorporating strategies that foster positive linguistic elements associated with social presence. This could involve engaging in collaborative discussions and creating an environment that encourages positive emotions and open communication. Further research could explore the relationship between linguistic elements and other types of presence, such as cognitive and teaching presence, in different communication contexts.

Keywords: linguistic inquiry and word count; LIWC; social presence; word count; Whatsapp.

1. INTRODUCTION

In the higher education context, computer-mediated communication (CMC) is omnipresent, and it is a part of the learning mode. The use of CMC in higher education is not only limited to the use of computers, but now it has also transcended to mobile applications. Participants usually exchange words or ideas by typing them into a web or mobile application. Studies have shown that different social media platforms have led to different varieties of language patterns (Mercy & Christa, 2020). One of the common mobile applications used in higher education is WhatsApp. It is an instant messaging application that facilitates real-time communication among individual users and organisations. This social application has attracted great interest in recent years owing to its capacity to minimise the physical barrier between people worldwide. Since WhatsApp is a social application, this study is interested in covering one crucial aspect of online learning: social presence.

The concept of social presence holds significant importance within the context of online learning. Social presence refers to the sense of being connected to other people when people are engaging with other people online. This engagement is significant for establishing meaningful relationships in mediated communication. When students and teachers have established social presence, it can enhance students' engagement and motivation to participate in the learning process actively. Students who experience a sense of social presence will likely remain in a course and achieve academic success (Boston et al., 2009; Liu et al., 2009).

Besides, linguistic analysis can be used to understand how social presence is represented in online learning. Pennebaker et al. (2014) stated that when a person chooses words in their speech or writing, it can provide valuable insight into their thoughts, emotions, and overall cognitive processes. Researchers can better understand how learners form social connections and collaborative activities by observing linguistic aspects of online interactions, such as word choice, the use of pronouns, reference words, and language style. Other linguistic indicators that researchers can use to identify social presence include the use of inclusive pronouns (e.g., "we" vs "I"), expressions of empathy or support, and the existence of conversational characteristics such as turn-taking and backchannelling. These linguistic features are useful to show evidence of social presence in a communication medium.

LIWC has been utilised in various contexts to analyse psychological attributes and cognitive processes, including higher education. For example, it has been employed to analyse students' essays and reviews by Yelp reviewers (Kovács & Kleinbaum, 2019). LIWC has also been applied to texts to identify linguistic markers of grandiose narcissism (Holtzman et al., 2019) and assess the language associated with cardiovascular risk (Andy et al., 2021). In addition, LIWC has been used to measure story sentiment, specifically the ratio of negatively to positively balanced words, in analysing scepticism about vaccine safety (Motta & Stecula, 2021). However, few studies use computer-aided software such as LIWC to analyse social presence in WhatsApp. Therefore, we intend to analyse the occurrence of social presence on WhatsApp applications using LIWC analysis. It is hoped that LIWC can provide some insights into the linguistic features and characteristics of language used in the selected context.

2. METHODOLOGY

This quantitative study utilised computer-aided software to analyse social presence. The LIWC analysis examined WhatsApp transcripts of 50 psycholinguistic attributes, including effect, cognition and perception, interpersonal focus, and social categories (Tausczik & Pennebaker, 2010). LIWC analyses transcripts word-by-word, comparing them with a dictionary related to various linguistic, affective, cognitive, and social processes. We can explore how social elements are constructed by analysing the language used in WhatsApp discussions through LIWC analysis.

The use of LIWC in this research paper is appropriate for several reasons. Firstly, LIWC has been established as a reliable tool for text analysis by numerous studies. Secondly, LIWC analysis allow researchers to explore various aspects of human communication that might not be detected otherwise using traditional qualitative or quantitative methods. For this study, LIWC2015 was used to analyse the WhatsApp conversation. To determine the linguistic features related to social presence, the researchers refer to the seven linguistic categories suggested by Zhu et al. (2019): clout, authentic, tone, first pronoun (I, We), positive emotions, negative emotions, and social processes.

The interpretation of the data was based on Pennebaker et al. (2015) and the LIWC2015 operator's manual. For the first category, clout, high score displays confidence and high expertise among the participants, but low score means tentativeness, humbleness, or even anxiety in speech. Next, the second category, authenticity with high score implies that the participants are honest, more personal and open in their conversation. But low score shows guarded, cautious, and distanced form of discourse. The third category is tone. Higher scores show positive expressions and a more upbeat conversational style, and low scores mean negative expressions with elements such as anxiety, sad or hostile. If the scores hover around 50, it has a level of ambivalence or mixed from both attributes. Meanwhile, the other four categories (first pronouns, positive emotions, negative emotions, and social processes) were interpreted according to the number of occurrences in percentages. The findings will be presented in the next section.

3. FINDINGS

Table 1 displays the description of the participants involved in this study. Table 2 depicts the results of the LIWC analysis based on seven chosen linguistic elements conducted on four different groups of postgraduate students' conversations on WhatsApp. The table includes information about the group's word count and the percentage of different linguistic categories identified by LIWC analysis.

Table 1: The Description of The Participants

Group	Students per group	Gender	Relationship history
A	3	Three females	Two members have known each other since they were undergraduates (~5 years), and two members were coworkers (~2 years). One member has a mutual relationship with two members.
B	4	Four females	Two members had known each other since they graduated (around 1.5 years ago), while the other two were new.
C	4	One male, three females	All of them are introduced during the course.
D	4	Two males and two females	They have known each other since beginning postgraduate study (~1.5 years).

Table 2: Results of The LIWC Analysis

Group	Word Count	Clout	Authentic	Tone	First Pronouns		Positive Emotions	Negative Emotions	Social Processes
					I	We			
Group A	2530	92.48	48.87	93.36	3.32	3.99	5.18	0.87	13.56
Group B	1850	94.02	27.76	99.00	4.00	3.46	7.41	1.30	14.11
Group C	4411	91.19	35.08	99.00	3.58	2.99	8.98	0.75	13.35
Group D	6655	90.38	31.76	98.17	3.44	4.04	6.61	1.11	14.06
Mean	3861.5	92.0175	35.8675	97.3825	3.585	3.62	7.045	1.0075	13.77

Group A had a total word count of 17 weeks of conversation, which was 2530 words. The group showed a high level of clout (92.48%), indicating a high level of confidence and showed their expertise in their messages. They used a balance of guarded and honest form of discourse (authenticity), which was 48.87% and the conversational tone of Group A was upbeat and positive most of the time (93.36%). Group A used the first-person plural pronoun "We" more than the first-person singular pronoun "I", 3.99% and 3.32% respectively. About 5.18% of Group A's messages contained positive emotions, while 0.87% contained negative emotions. Social processes made up about 13.56% of Group A's messages.

Next, the total number of word count for Group B was 1850 words. This group showed a high level of clout in their messages, with a score of 94.02%. However, Group B used a more distanced form of discourse (authenticity), which was 27.76%. The group showed a very high level of tone (99.00%), indicating an upbeat and positive style of conversation. Group B used the first-person plural pronoun "I" (4.00%) more than the first-person singular pronoun "we" (3.46%). About 7.41% of Group B's messages contained positive emotions, while 1.30% contained negative emotions. They also used a high percentage of social processes (14.11%) in their messages.

For Group C, the word count for this group was 4411 words. This group showed high levels of clout in their messages (91.19%). However, Group C used a more distanced form of discourse (35.08%) in their messages. The group showed a high level of tone (99.00%). Group C used the first-person plural pronoun "I" with 3.58% more than the first-person singular pronoun "we" with 2.99%. About 8.98% of Group C messages contained positive emotions, while less than 0.75% contained negative emotions. They also used a high percentage of social processes (13.35%) in their messages.

In the last group, Group D, the word count was 6655 words. The group showed a high level of clout (90.38%). They used more distanced discourse (31.76%) and a high level of tone (98.17%) in their messages. Group D used the first-person plural pronoun "we" (4.04%) more than the first-person singular pronoun "I" (3.44%). About 6.61% of words showed positive emotions, and 1.11% of negative emotions. They also used a high percentage of social processes (14.06%).

Overall, the LIWC analysis shows that all four groups of postgraduate students used WhatsApp to communicate effectively. They showed a high level of social presence by using social processes and first-person plural pronouns in their messages. They also showed a positive and upbeat conversation style, with a low level of negative emotions. However, some groups used a distanced form of discourse, indicating a lack of authenticity in their messages. The discussion of the findings will be explained in the next section.

3.1 Discussion

Language patterns on WhatsApp can inform social presence to a certain extent. From the findings, generally, postgraduate students were able to show social presence based on the linguistic features presented. In this study, all students could show a high “clout” level (i.e. confidence or high expertise) in the discussion. As most students are working adults pursuing postgraduate studies, they already manifested a certain level of expertise in their academic background. When they have the confidence to talk about the content they know, it is much easier to engage with other students and eventually develop social presence. Their experience and educational background contribute to demonstrating “clout”.

However, despite their high confidence level and expertise in the LIWC analysis, they appeared to be quite guarded in the discussion. Most of the participants communicated in a cautious and distant manner. Even though WhatsApp is commonly used for more personal and close communication, the students showed a significant distance in their conversations. One of the reasons is that the instructor is in the WhatsApp group, and the intention to save face is present. As Asian culture is collectivistic, hierarchical, and face-conscious (Rudowicz & Ng, 2003), the students displayed a considerate and cautious attitude by clearly being mindful of their choice of words. Furthermore, these students are still bound in the academic domain, which is an institutionalised context where certain conversational rules are applied. This domain is also associated with formal and serious discourse since it requires achieving academic goals instead of socialising.

In terms of tone, positive emotions and negative emotions, all groups demonstrated upbeat and positive expressions and displayed more positive emotions than negative emotions in the discussion. The positive group dynamics and good attitude from the group members can contribute to such tone in the overall discussion. As this study is conducted in a fully online context and most of them just met virtually, keeping a positive environment is necessary. That is one possible reason positive emotions are displayed more than negative emotions. The tone mentioned here may also be influenced by cultural norms, particularly in Asian cultures where social order and harmony are highly valued, and conflicts should be avoided (Rudowicz & Ng, 2003). Thus, creating a positive climate contributes to social presence in the learning environment.

In addition, research has also shown that WhatsApp groups can create a stronger sense of community (Suardika et al., 2020). One indicator is using the first pronouns "I" and "We". Based on our thorough analysis using concordance software, "we" usually involves the group action to do something. For example, the phrase "we need," such as "we need to decide the content, we need to read the article, and we need to finalise some elements," was quite common in the WhatsApp discussion. The "we" pronoun indicates a group call for collaboration and a sense of belonging. It also describes how individuals connect themselves to others (Zengilowski et al., 2023). Meanwhile, the use of "I" was prevalent when stating one's "okay" and "fine" condition, such as "I'm okay, I'm fine", and stating their emotions, such as "I'm nervous, I'm worried, I'm a little bit confused". The pronoun "I" showed that these students communicated openly and could express emotions freely on WhatsApp. By observing the use of first pronoun and additional concordance analysis, we could see how students can portray social presence on this platform.

The use of social processes category such as "his, mother, they" was also observed in LIWC analysis. The number was not considered high, meaning they avoided including words related to family, friends, and male/female references in the discussion. The discourse where the discussion occurs is academic, which can be a reason for the lack of social processes category. They may prefer to use names or neutral references like "you" when addressing their friends in the discussion.

Using LIWC analysis as the only analysis in the study may serve as one of the study's limitations. LIWC can be used in complement with qualitative or corpus analysis for rich data interpretation. Researchers can provide examples from the source text to give meaning to the data analysed. Furthermore, it can only process text, not emojis, stickers and others. Despite limitations, LIWC is very useful as a preliminary study of more extensive research, especially for data-driven research. The findings can be used as a preliminary result where the researcher can preview or have a broad idea of what the data can offer regarding linguistic elements that can be used to identify social presence in a short time.

4. CONCLUSION

Overall, the LIWC analysis showed that all four groups of postgraduate students showed high social presence through a high level of clout, tone, positive emotions, use of the first pronoun "we" and some social processes in their messages. They also showed a positive and upbeat conversation style, with low negative emotions. However, some groups used a distanced form of discourse. This study highlights that linguistic elements can imply the existence of social presence in WhatsApp to a certain extent. This study shows how educators and instructional designers can improve the online learning experience by incorporating strategies that enhance social presence through linguistic elements. In addition, providing a learning environment that promotes social presence can encourage positive emotions and open communication among students. Further research could explore the relationship between linguistic elements in LIWC and other types of presence, such as cognitive and teaching presence in different communication contexts.

REFERENCES

Andy, A.U., Guntuku, S.C., Adusumalli, S., Asch, D.A., Groeneveld, P.W., Ungar, L.H., & Merchant, R.M. (2021). Predicting Cardiovascular Risk Using Social Media Data: Performance Evaluation of Machine-learning Models. *JMIR Cardio, 1*(5), e24473. https://doi.org/10.2196/24473

Boston, W., Diaz, S., Gibson, A., Ice, P., Richardson, J. C., & Swan, K. (2019). An exploration of the relationship between indicators of the community of inquiry framework and retention in online programs. *Online Learning, 13*(3). https://doi.org/10.24059/olj.v13i3.1657

Holtzman, N. S., Tackman, A. M., Carey, A. L., Brucks, M. S., P. Küfner, A. C., Deters, F. G., Back, M. D., Donnellan, M. B., Pennebaker, J. W., Sherman, R. A., & Mehl, M. R. (2019). Linguistic Markers of Grandiose Narcissism: A LIWC Analysis of 15 Samples. *Journal of Language and Social Psychology*, *38*(5–6), 773–786. https://doi.org/10.1177/0261927X19871084

Kovacs, B., & Kleinbaum, A. M. (2020). Language-Style Similarity and Social Networks. *Psychological Science*, *31*(2), 202–213. https://doi.org/10.1177/0956797619894557

Liu, S. Y., Gomez, J., & Yen, C.-J. (2009). Community college online course retention and final grade: Predictability of social presence. *Journal of Interactive Online Learning, 8*(2), 165-182.

Motta, M., & Stecula, D. (2021). Quantifying the effect of Wakefield et al. (1998) on scepticism about MMR vaccine safety in the U.S. *PLOS ONE*, *16*(8), e0256395. https://doi.org/10.1371/journal.pone.0256395

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLOS ONE, 9*(12), e115844. https://doi.org/10.1371/journal.pone.0115844

Pennebaker, J. W., Boyd, R., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. University of Texas at Austin. https://doi.org/10.15781/T29G6Z

Rudowicz, E., & Ng, T. (2003). On Ng's Why Asians Are Less Creative Than Westerners. *Creativity Research Journal, 15*(2-3), 301-302. https://doi.org/10.1080/10400419.2003.9651423.

Suardika, I. K., Suhartini, L., & Pasassung, N. (2020). Using WhatsApp for Teaching a Course on the Education Profession: Presence, Community and Learning. *International Journal of Mobile and Blended Learning, 12*(1), 17-32.

Tausczik, Y.R. and Pennebaker, J.W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology, 29*(1), 24-54. http://dx.doi.org/10.1177/0261927X09351676

Zengilowski, A., Lee, J., Gaines, R. E., Park, H., Choi, E., & Schallert, D. L. (2023). The collective classroom “we”: The role of students’ sense of belonging on their affective, cognitive, and discourse experiences of online and face-to-face discussions. *Linguistics and Education, February 2023,* (73), 101-142. https://doi.org/10.1016/j.linged.2022.101142