Emoji Usage by Android Consumers: A Gender-Based Inquiry

Muhammad Shahbaz[1] Tabassum Razzaq[2]

Muhammad Riaz Gohar[3]

Abstract

In recent years, android apps have achieved unbelievable popularity and turn out to be a new source for Computer-Mediated Communication (CMC). There is gender-based imbalance of language use. This paper specifically aims to explore the gender-based differences in the use of Emoji. Several research studies have been done to identify behaviour of male and female toward emoji usage. Gender-specific studies are always meaningful for understanding how much males and females vary in their communication styles. However, until now not much is known that to what extant both genders have difference in the use of emoji. To fill such a gap, this study attempts to explore the use of Emoji in conversation through gender lens. The analysis depends on the data collected from android app with the help of Kika keyboard. For this purpose, 144,478 users from 180 countries, and their 510 million messages are selected as data set.A multi-dimensional statistical analysis has been conducted from many view pointsin the use of Emoji, taking account of the differencesand preferences. The findings demonstrate that the use of emoji can noticeably differ between males and females. Consequently, various implications are purposed that can help in creating better understanding about gender-based use of emoji and for keyboard layout developers.

Keywords: Gender, Kikaboard, Consumerand Emoji

Introduction

[1] Assistant Professor, Department of English, GC Women University Sialkot
[2] M. Phil Scholar, Department of English, GC Women University Sialkot
[3] Ph. D Scholar, Department of English, University of Gujrat, Gujrat

Every researcher has a purpose behind choosing an area for research work or investigation. For carrying out this study, researcher has a personal reason to choose this area. The researchers study about the gender-based differences in the use of verbal and non-verbal communication and started to explore the non-verbal differences in conversations. This work is an effort to provide answers aboutgender-based use of emoji, preferences and their frequency.

Emoji have gained astonishing popularity in recent years and are interwoven in our daily communication. Text messages and pictures are considered typical traditional representation of information but emoji are taken as more communicative, realistic and rich in meaning so more preferred by smart-phone users. Some people prefer to use emoji to share information instead of typing a text, for instance, on April 11, 2015, Andy Murry, announced his wedding on twitter by using 51 emoji. Some current research have paid attention to understand the use of emoji across app and cultures but no significant effort is made to understand the emoji usage behaviours across genders. Like face to face

conversational differences, gender-based differences also exist in non-verbal cues. When it comes to the study of gender-based non-verbal differences, Some researchers claim that very little is known whereas other argue that there is enough work relating to gender-based differences but not systematically based on sociolinguistic variables. This study seeks out some of this gap in previous research. This study adds to the field of sociolinguistics as well as pragmatics and the focus is on usage of emoji by women and men in text conversation.

Numerous studies relating to gender speech styles have been conducted using real life conversations (Coates, 1986; Tannen, 1991; Githens, 1991), but little has been done in other frameworks through different methods. Several recent works exposed that language and conversational patterns from online source are useful to understand and explore gender-based conversational differencesin individuals and populations. This effort includes utilizing android apps of communication to understand the use of emoji related to gender-based conversation. With the growing popularity of communication messenger, many researchers have shown interest in exploring

computer mediated communication. However,little research is based on the usage ofinternet conversation from gender perspective. Motivated by the preceding literature about gender differences, android apps are consideredas new communication messenger and researcher investigates the use of emoji in mobile conversation.

Study of language and gender is an important field and one of the hot issues in this field is how male and female use their communication skill to interact with one another. It has so far been observed that men and women have many differences at every level of the language organization. There is no doubt that men and women are different biologically as well as in many other aspects of life.No doubt, the investigation about this copious and ancient issue of language and gender is triggered by the hypothesis that women language is powerless language and this supposition challenged the entire concept of 'gender' by underlining the performance of male and female.

Language usage is influenced by many factors such as setting, audience, and even the gender of both the speaker and listener.

However, the most influential factor is society. The men and women have many differences in their communication strategies however, in this paper, the main focus is on the gender-based use ofemoji. The present study contributes in the existing literature of Language and Gender and Feminist Linguistics. This work is unique from two perspectives; firstly, the data set is largest, from a renowned app named Kika. The data set encompasses180 countries (144479 consumers)and their messages (510 million)assembled in90 days that make the result more comprehensive. Secondly, various descriptive and inferential statistical analyses helped in understanding different facets of emoji usage.

Following findings are made:Gender-based differences are found in frequency of emoji. Female users use one emoji or more than one emoji in one message, in contrast, maleuse multiple emojis.Both genders have different preferences; female use face related emojis whereas male prefer heart related emojis.

Literature Review

This section starts with brief discussion of the related background and literature. With the rapid use of emojis, various researchers have shown curiosity to know more about the usage of emoji but very little is known about emoji usage from gender perspective. The researchers investigate the impact of gender on emoji usage in text conversation.

Emotions, emoticons and Emojis

Facial expressions and body language are part of our daily communication to enhance effects of our sentiments and emotions. In 93% of everyday communication, people use symbolic representation of their messages and emotions. Emoji is a set of pictographic Unicode characters that use to express humour, irony, and sarcasms and intimacy. Emojis were created and introduced in Japan for the first time in late 90s.All common mobile platforms now support emojis. Emoji were first specified by Unicode standard in version 6.0. 722 emoji are standardized in version 6.0, by Unicode in 2010.

Tossell et al. (2012) investigated how emoticons are exercised, in text messaging and in particular, how much gender differences exist in the frequency and variety of emoticons used. For analysis, data from twenty-one smart phones was taken for last six months. He observed that females are more likely to use emoticons where as males use a more distinct range of emotions. Hwang (2014) reported from data set of 126 students from Korea, female students used emoticons to express feeling and intimacy than male. For them, emoticons played a role as a facilitator. Wolf (2000) explored emoticon practices among different gender groups and concluded that there were significant differences among same and different gender groups.

Emoticons have limited morphological variation of symbols, non-standard formation and their use bring in some questions to data analysis where as emoji has standard creation. They not only express emotion but also able to represent other objects as well. There exist prior research on emoji usage across culture and boundaries but there exist limited work on emoji usage from gender perspectives. Some prior studies reported the impacts of gender on emojis. For instance,

Nishimura (2015) examines sample that consists of200 participants from a Japanese blog site established in 2004, having 121 major categories and some 7000 sub-categories. He finds that women are more inclined to send emojis compared to men. Similarly, however, we could not find any study with largest sample about emoji usage from gender perspectives except Chen et al (2017). To overcome this research gap, the researchers explore the gender-based use of emoji on large-scale through Android users.

Lu et. al (2016) study the data set consisting of 400 emoji containing messages of 3.87 million active users from 212 countries and find gender of user can be predicted through the emoji usage. Chen et.al (2017) efforts to explore the gender-based use of emoji, for these purpose 134419 active participants from 183 countries are selected and 401 million messages are saved in 90 days. The findings show that there is a difference in emoji usage between both genders.

Kika keyboard

Kika keyboard is a major Google Play app for input. In 2015, it was ranked as one of the top downloaded apps for Google play store. Kika assists 82 languages and has input of 1281 emojis.For research purposes, Kika collects the necessary data, e.g. the language, country information and content of the text. Kika records no personal or unethical data from the user input.

Research Methodology

This section provides overview of research methodology, approach, dependent and independent variables, population, specified group, sample, tool of research, procedure of data analyses and ethical considerations for this research.

Data collection and description

Unlike the other studies about non-verbal differences, this paper makes use of the sample gathered through Kika keyboard that is an input app in Google play. It secured millions of downloads across the globe and support 82 languages having more than 3000 emojis and emoticon. The researchers use version 4.0 to search the data set.
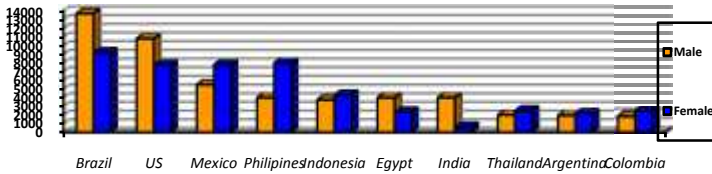
Figure 1: Top 10 Countries with in Action Consumers

Procedure of data analyses

Information is collected from only those users who are agreed to Kika's Privacy Policy that no secret codes or hypersensitive data are noted down. All the information has been anonym zed before the examination. As an input technique, Kika is not limited to only one application such as twitter; it can attain the contexts of different apps where allowed and introduced. Having this feature, Kika enable the researcher to have thorough study of emoji usage. The sample consist of

Shahbaz et al.

144, 479 participants from 180 countries, and 510 million text-messages from November 4, 2017 to January18, 2018. An anonym zed consumer identifier will be used to signify every active user who communicated through Kika keyboard (at least one message). Demographic information of every participant is given and the number of messages typed by him/her. From there the researchers report the demographic division ratio of both genders in data-set. Data from top ten world regions with highest proportion of usage is(Brazil, US, Mexico, Philippines, Indonesia, Egypt, India, Argentina and Colombia) presented in Figure 1. Almost seventy-one percent of overall consumers are from these 10 countries. However, as a whole, male and female distribution is not equal. As illustrated in Figure2, 46.5% are male and 53.5% are female.
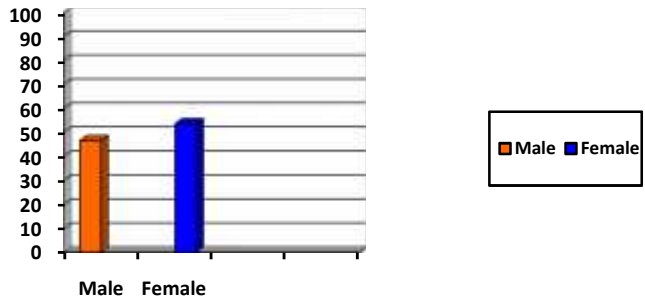
Figure 2: Gender Distribution

Ethical Consideration

As a matter of fact, this research enclosed sensitive personal facts of users; hence cautious measures are taken to uphold the ethics in this research as it is a critical concern for the researchers. Kika exchanges the device ID of each participant and it is made sure that no sensitive data is saved and all the data through which a participant can be identified had been anonymized before the analysis. Data is securely saved (HIPPA-complaint cloud server) that is strictly sanctioned by the organization. For ethical considerations, all words are removed and only emojis are extracted for analysis. The message length was calculated and emojis were labeled with sentiment by studding the

contexts. Kika employees constantly governed the entire procedure to make sure the agreement with the public secrecy policy acknowledged by Kika. The researchers tried hard to provide original, authentic knowledge and truth without any kind of error. It was verified on the part of researchers that there is no forgery, fabrication and misrepresentation of any finding which could alter a real scenario.

Limitations

Like other studies, this study also has some limitations which can potentially affect the analysis and derived results. The data set of this study only includes the participants who freely reveal the gender information and the user allocation is not consistent between all countries, consequently, there could be some prejudice. Furthermore, this work is not based on gender differences among limited or controlled user group because of large and bulky sample size. To the best possible knowledge of researchers, sample of this study is the largest. The larger data set of this study makes the analysis broader and the results significant.

Data Presentation and Analyses

This empirical study means to search whether there are some gender-based differences in emoji usage and which emoji is used most by males and females. This study mainly focuses on the Usage Pattern Analysis. Emojis are considered semantically rich even without words. For this purpose, macro level descriptive analysis is done by making use of two-tailed z -test. This research focuses on usage patterns analysis.

Usage pattern analysis

The researchers focus on gender-based analysis of emoji usage pattern from several aspects, including frequency, choice, and consecutive/discrete pattern and frequently use emojis.From the sample, researchers examine 31 million messages comprising minimum one emoji, crediting for 7.5% of all messages. Beyond the extensive use of emojis, the researchers are probing whether there is any difference in emoji usage among males and females. How often do male and female

Shahbaz et al.

participants utilize emojis in messages? Which emoji is more frequently used by male and female users, respectively?

Frequency of emojis in Messages

Previous studies (Ablon et.al, 2013; Marianne & Mahzarin, 1992; Briton & Hall, 1995) show that female are more non-verbal expressive than male. Prior studies (Dresner & Herring, 2010; Daantje & Jasper, 2007) have also noticed the gender differences in the frequency of emojiusage. Emojis are also considered as non-verbal cue, the researchers are interested to know whether a similar difference is there in emoji usage frequency. The researchers begin with exploring which gender frequently uses emojis in their messages.

Table 1: Gender-based occurrence of emoji (**'': p-value <0.01)

Gender	% of Emoji in total	Emoji% (2017-11)	Emoji% (2017-12)	Emoji% (2018-1)
Male	7.03	6.9	7.00	7.50
Female	8.00	8.0	7.80	8.5
z-score	-348.48**	-200.83**	-216.97**	-183.02**

Table 2: Emoji usage pattern

	Male %	Female %	z-score
Pat 1	40.00	41.00	-60.81
Pat 2	57.85	56.00	78.00
Pat 2.1	38.00	37.00	56.01
Pat 3	10.07	10.00	-34.98

The gender-based percentage of messages containing emoji is compared as whole and monthly. As displayed in Table 1, the 7.03% of messages post by male participants consist of minimum one emoji; where as 7.96% of female messages contain at least one emoji. A significant difference is found under two tail z-test so it represents that females have higher percentage in using emojis.

Finding (F1): In general, the analysis of data set reflected that emojis are more frequently used by female than male.

Consecutive/Discrete Usage Patterns

We have seen the frequency of emoji usage in the above section. Some messages have only one emoji while others have more than one emoji. Now the researchers want to discover whether both genders have different input patterns for emoji.

Typically, there are three input patterns.

Pattern (Pat1): Message having one emoji, e.g., I am happy.

Pattern (Pat2): Use of more than one emojis in one message, e.g.

I am happy.

Pattern (Pat3): Posting of multiple emojis in one message, e.g., I am happy.

In pattern 2, the use of this type of pattern is an emphasis to show stronger impact. Here the proportion of these three patterns are identified and counted then by performing z-test, gender difference in usage pattern is measured. Bonferroni correction is applied to regulate p-values for multi-hypothesis testing. The results are succinctly summarized below. Interesting gender-based differences are shown in Table 2. Female participants utilize single emoji or more than one

emojis in a message whereas male participants' tend to use consecutively multiple emojis in one message.

Finding (F2):

Both genders use emoji patterns differently. Females utilize one emoji or discretely use more than emoji in single message. On the other hand, male use multiple emojis consecutively.

Frequently Used Emojis

We have found the gender-based frequency of emoji usage. Yet little is known about which emojis are more likely to use. Do men and women use same emojis? To find out the answer, favourite emojis used by both genders are listed. For all emoji, the percentage of its occurrence is counted in male and female messages. Ten most commonly use emojis are listed in figure 3. No significant difference is found in the usages of these ten emojis, first five have same frequency of usages by both genders.
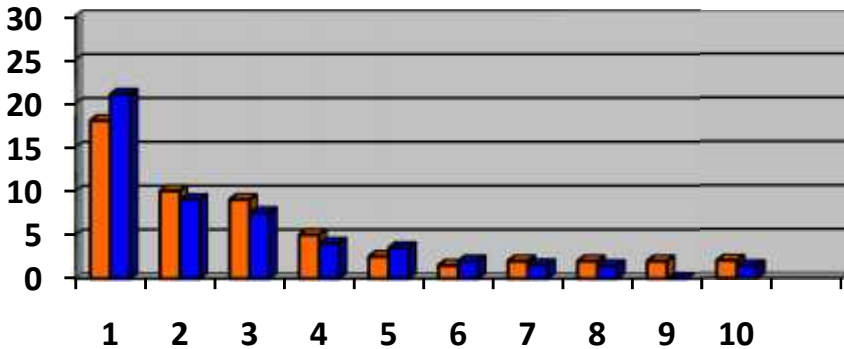
Figure 3: top ten frequently typed emojis

However, there is difference for particular emojis, emoji with tears of joy is most popular emoji among male and female but its usage is found 18% in males' messages where as in female 22% with a difference of 4%.

Finding (F3): both genders impart the same top 5 most used emojis. However, there is a sharp difference in usage proportion of some emojis.

Summary of Findings

Until now, the researchers have explored the usage behaviours for emojisand reported several gender-exclusive differences. From the

19

findings F1-F3derive from our descriptive analysis, basic understanding is developed how emojis are used by females and males. Statistically, there is quite significant difference in behaviour of female and male users in using emojis. These findings confirm the initial hypothesis of this study.

Implications

Following the previous findings, some implications are given;As there are found some gender -specific differences in the use of emojis, the most practical implication in the field of software development for software developers is to provide some hints of problems while searching right or left for emojis so they may design new features in keyboard layoutto improve the users experience on smart phone keyboard. Keyboards' layout should be design in such a way to make users aware about gender. Based on our interpretations, keyboard developers can improve situation-related and temporal-related emoji usage patterns, recommendations under some specific contexts.

This research work enlightens some observations and findings to distinguish users' gender through their emoji usage that will be useful in daily life. This study adds to field of socio-linguistics.

Above all, this study links with the key subject matter of feminism. In addition to that, this thesis will incorporate much of the work to the field which will be useful for future researchers in perspective of gender studies.

Limitation and Future work.

Like every empirical study, this study has its own potential threats and limitations that can influence the overview of results. One major threat may arise in data collection. For this purpose only, active users in sample are considered by the Kika keyboard whereas there are many other input methods. Hopefully, such a threat could not affect the results, as the sample size encloses a huge mass of real-world users from several countries that make this study statistically comprehensive. However, in addition to gender, there are many other perplexing user-profile factors that may affect, suchas country, cultureand language.

Due to time and space restrictions, this study is based on only usage pattern analysis and analyses from other aspects are excluded, for example CMC scenario and sentiment analysis. The researchers suggest the future researchers to integrate such information along with gender that make up the gender difference more visible.

Conclusion and future Work

This work is based on gender-based empirical study of emoji usage by making use of large sample size gathered through Kika.The data set consistof510 million messages posted by 144,479participants from 180countries in duration of three months. Data set comprise of 46.5% of males and 53.5% females. A multi-dimensional statistical analysis is conduct from several aspects, including the frequency, input patterns and preference. Various statistical tests are employed to verify the findings. In actual fact, various differences are reported and gender differences are found in the usage of emoji. On the basis of observations and findings some implications are introduced, such as distinguishing the gender of users through their emoji usage patterns.

To the best of our knowledge, the sample of this study is the largest to

analyse emoji usage through a gender lens.

References:

Ablon, S. L., Brown, D. P., Khantzian, E. J., & Mack, J. E. (Eds.). (2013). *Human feelings: Explorations in affect development and meaning*. Routledge.

Briton, N. J., & Hall, J. A. (1995). Beliefs about female and male nonverbal communication. *Sex Roles*, *32*(1-2), 79-90.

Chen, Z., Lu, X., Shen, S., Ai, W., Liu, X., & Mei, Q. (2017). Through a Gender Lens: An Empirical Study of Emoji Usage over Large-Scale Android Users. *arXiv preprint arXiv:1705.05546*.

Derks, D., Bos, A. E. R., & von Grumbkow, J. (2008). Emoticons and online message interpretation. Social Science Computer Review, 26(3), 379-388.

Dresner, E., & Herring, S. C. (2010). Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication theory*, *20*(3), 249-268.

Hwang, H. S. (2014). Gender differences in emoticon use on mobile text messaging: evidence from a Korean sample. *International Journal of Journalism & Mass Communication*, *2014*.

LaFrance, M., &Banaji, M. (1992). Toward a reconsideration of the gender-emotion relationship. *Emotion and social behaviour*, *14*, 178-201.

Lu, X., Ai, W., Liu, X., Li, Q., Wang, N., Huang, G., & Mei, Q. (2016, September). Learning from the ubiquitous language: an empirical analysis of emoji usage of smartphone users. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 770-780). ACM.

Nishimura, Y. (2016). A sociolinguistic analysis of emoticon usage in Japanese blogs: Variation by age, gender, and topic. *AoIR Selected Papers of Internet Research*, *5*.

Pohl, H., Domin, C., &Rohs, M. (2017). Beyond Just Text: Semantic Emoji Similarity Modeling to Support Expressive Communication. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *24*(1), 6.

Tossell, C. C., Kortum, P., Shepard, C.,Barg-Walkow, L. H., Rahmati, A., &Zhong, L. (2012).A longitudinal study of emoticon use in text messaging from smartphones. *Computers in Human Behaviour*, *28*(2), 659-663.