

Moving Ethnography Online: Analysing the Social Meaning of Manglish in Youth WhatsApp Chats

Nur Husna Serip Mohamad, Nurhilleny Rosly, Nur Liyana
Johari, Nurul Syafieqah Jaafar and Ahmad Aminuddin Soopar

Akademi Pengajian Bahasa, Universiti Teknologi MARA, Malaysia

Email: nurhilleny@uitm.edu.my, liyanajohari@uitm.edu.my, nsyafieqahj@uitm.edu.my,
aminuddins@uitm.edu.my

Corresponding Authors Email: nurhusna@uitm.edu.my

To Link this Article: <http://dx.doi.org/10.6007/IJARBSS/v12-i8/14410>

DOI:10.6007/IJARBSS/v12-i8/14410

Published Date: 19 August 2022

Abstract

This article contributes to the changing landscape of fieldwork practices within sociolinguistic research, specifically the practice of online ethnography. It makes a case for the significance of online chats in observing the construction of ethnic identities and social intimacy among young Malaysian English (henceforth Manglish) speakers. The rise of Manglish discourse has been hardly ignored in digital settings. Thus, the rationale for choosing Manglish as a language of communication within online spheres such as Instant Messaging is unclear. This article argues that the socially structured meanings that underlie Manglish practices deserve exploration as it would benefit from online ethnography. Drawing on Androutsopoulos's (2008) discourse on online ethnographic approach, this paper shows how researchers could use WhatsApp chats to explore the linguistic and social behaviour of Manglish speakers. The findings suggest that the interrelation of speakers' online behaviour and linguistic choice can denote their ethnicity, in-groupness, stances, and regional identity. Online ethnography is therefore significant and should not be limited to linguistic sources but also used as a mode of participating in, as well as observing, the construction of identity that reflects the 'offline' aspect of speakers' social lives.

Keywords: Identity, Manglish, Online Ethnography, Social Meaning, Whatsapp, Ethnicity

Introduction

Research on social identities in the third wave of variation studies has mainly employed ethnographic methods to seek out linguistics' relation with social practices and how it contributes to the construction of identity. The ethnographic approach has been applied to studies of speech style-shifts and variation among different groups of speakers (Moore & Podesva, 2009; Moore, 2003). For instance, in their study of tag questions, Moore and

Podesva (2009) not only observe linguistic sources such as realisation of /t/ in tag questions but also other semiotic signals of social differences such as habits, appearances, and clothes. These observations allow them to draw inferences about the speakers' social group memberships. The proliferation of digitalised communication has prompted the relocation of ethnographic frameworks into Computer-mediated Communication (CMC) settings (Androutsopoulos & Beiswenger, 2008; Androutsopoulos, 2008). In other words, research into online linguistic practices has opened up interesting avenues of investigation, including in-group identity and performative construction of an 'online' communicative identity among the speakers (see for instance Androutsopoulos, 2015; Bamman et al., 2014; Kytölä & Androutsopoulos, 2012; Shrooten, 2012; Tsiplakou, 2009).

For CMC research, quantitative paradigms can demonstrate the trajectory of linguistic patterns among a group of speakers (Androutsopoulos, 2014). However, current third wave studies emphasise the importance of micro-level interactions to understand the myriad ways that speakers derive meaning from their speech style. This includes any linguistic material that serves an expressive purpose (Eckert, 2019). As Kytölä and Androutsopoulos (2012) put it, 'corpus-analytic and quantitative approaches may overlook the pragmatics of language use in micro-level interaction and even leave it unexplained or under-contextualised'. In other words, interaction in local context is key to understanding speakers' speech styles and this paper shows how online ethnographic analysis could add more insights in understanding dialects in social and interactional contexts.

However, debates have emerged surrounding the validity of online ethnographic methods that it contributes to inaccuracy, inability to interpret gestures, threaten authenticity and so on (see for instance Abidin & De Seta, 2020; Boellstorff, 2015; Bengtsson, 2014; Gobo, 2008; Illingworth, 2001). In her review of the field, Bengtsson (2014) addresses the challenges faced by digital ethnographers and ethical dilemmas when conducting ethnographic work. Bengtsson suggests that conducting online ethnography while the researcher is at home, living his or her own personal life, may reduce one's ability to understand a different culture comprehensive. Similarly, De Seta (2020) in her reflection on online ethnography, discusses the limitation and struggles of participating at 'site'. De Seta questions the meaning of being presence online and highlights ways to prove her active presence in the field. In another reviews on online ethnography, Abidin and De Seta (2020: 9) claim that online methods often cause 'anxieties, challenges, concerns, dilemmas, doubts, problems, tensions, and troubles' due to problems in handling the interactions and acquiring authentic conversations. Following these issues, this paper argues that online ethnography can successfully be conducted, where researchers can actively participate in the participants' life, beyond real-time interaction. In fact, meaningful conversations can be acquired easily if the real time interaction is not disrupted by third parties, in this context, the researchers themselves. Hence, the present study tackles online ethnographic issues by expanding on Androutsopoulos's (2008) approach to online ethnographies, in which a systematic observation can be conducted in an offline context, namely after the conversation has occurred. It is believed that this method allows the researcher to overcome dilemma between own and another culture, as well as reduce disruptions in acquiring real conversations.

Moreover, there are lack of methodological reflections on the field of online ethnography which motivates this paper. In this paper, we attempt to draw on the first author's (addressed as 'the researcher' throughout the paper) extensive social media fieldwork experiences in studying Manglish in WhatsApp Messenger and offer some thoughts on how to conduct qualitative research in the digital setting, especially in studying dialects. Specifically, this paper demonstrates that online ethnographic methods such as online observation, in-depth interview, and textual analysis can provide deep, localized knowledge of how certain Manglish features are used by the ethnically diverse participants. To do this, we show how the researcher adopted the Discourse Centred Online Ethnography (DCOE) (Androutsopoulos, 2008) approach into her data collection procedures and the way she positioned herself when interacting with the young participants. Such approach contributes to the current understanding of Manglish in local, social and interactional contexts, where Manglish features can be associated to in groupness, localness, or ethnic-ness. This paper also analyses extracts of conversation to explain how the first author arrived at her association of Manglish with the specified values.

Furthermore, this study underscores the relevance of the digital discourse by expanding upon previous Manglish work concerning computer-mediated communication conducted in different media forms of digital platform, such as Facebook and Twitter. Unlike previous studies that have mostly focused on the written interactions and functional aspects of Manglish features (see for instance Tay et al., 2016; Stapa & Shaari, 2013), the values this paper identify contribute substantially to the study of Manglish. Importantly, when employing an online ethnography, this paper shows that the analysis of WhatsApp chats is not restricted to textual information, but also addresses other multimodal aspects, such as symbols and emoticons.

In the following sections, we first provide an overview of online ethnography and how it emerged from the traditional ideas. We then defines Manglish and functional aspects of the features. In the methodological section, we demonstrate the researcher's own fieldwork experiences to discuss three major issues in online ethnography including the researcher's identity; initial observation of the data, during and post data collection.

Literature Review

This section discusses how online ethnography has become a distinct branch of ethnographic studies particularly due to the rise of CMC studies. The challenges of online ethnography are addressed to justify the need for a methodology known as Discourse Centred Online Ethnography (DCOE). The next section defines the term 'Manglish' and introduces its linguistic features as well as studies in various Malaysian CMC settings.

Online Ethnography

'Digital ethnography', or 'virtual ethnography', are ethnographic examinations of the social interaction of real-life cultures in virtual environments (Underberg & Zorn, 2013; Given, 2008). Grills (1999) argues that ethnographic studies can provide insights on expressions such as dilemmas, frustrations, and relationships. Though the expressions have become a problem for online ethnography, as it occurs in the virtual world of cyberspace, we argue that it is possible to conduct ethnographic studies virtually.

Within sociolinguistics, conventional CMC research is constantly challenged by digital communication's rigid social and pragmatic conditions. The lack of spoken language data and user socio-demographic information in virtual ethnography limits variation analysis. Androutsopoulos (2013) highlights one issue plaguing online data collection: the vast amount of digital data from distinct semiotic resources complicates the understanding of social contexts and meanings. However, as online ethnography research evolves, such views towards online ethnography should change with it. Greschke (2007) divides online ethnography into two main types. The first focuses on the Internet in daily undertakings, investigating how new communication technologies incorporated into the life and the culture of a community (Greschke, 2007). This is a blend of both online and offline ethnography, with offline activities relating receiving equal or more attention than their online counterparts. The second type concerns how everyday life takes place on the Internet, theorising the web as a medium for cultures to thrive. Previous ethnographic studies involving role-playing sites and online chat rooms have almost all prioritised participation observation but only few studies focused on areas entailing involvement and proximity with the participants (Doring, 2003). Hence, this paper illustrates the dynamic proximity that the author's experienced with her online subjects.

Androutsopoulos (2008), a pioneer of online ethnography, introduced the Discourse-Centred Online Ethnography (DCOE) approach to overcome issues pertaining to online participants. The two pillars of DCOE are systematic observation and contact with the subjects (Androutsopoulos, 2008). The first pillar focuses on insights into discourse practices and language patterns used, while the latter identifies a wide array of perspectives within a chosen field and the actors who exemplify different participation formats (Androutsopoulos, 2008). DCOE uses insights from ethnographical observations as the main setting for data selection, analysis, and interpretation to highlight relationships between the production and reception of digital texts. It is believed that DCOE can overcome the challenges of online data collection. Summarizing comprehensive observations and participant interactions illustrates more than just textual observation, as shown by other discourse analysis frameworks and language-centred CMC studies (Stapa & Shaari, 2013; Ling & Baron, 2007). Since the development of research in the field of CMC, new technologies have developed, leading to new modes of CMC. The development is best described by Androutsopoulos (2006) in his identification of the main 'waves' of CMC-related research. The first wave described the language used in cyberspace as distinct, homogeneous, and exclusive to members. In this phase, the description of online language was seen to have both written and spoken features while maintaining differences between asynchronous and synchronous modes. The second wave acknowledged technological interplay and social and contextual factors in shaping CMC practices, and the third wave heavily focused on the role of linguistic variability in the form of social interaction and social identities formed in cyberspace. On this note, this paper asserts that ethnographic studies have evolved in parallel with CMC studies and that DCOE has the ability to tackle issues pertaining to CMC and online ethnographic studies.

In his study on Facebook language practices, Androutsopoulos (2013) analyses the Facebook walls of a small group of Greek secondary school students. He also coins the term '*networked bilingualism*', which refers to multilingual practices bound by two interrelated processes: being networked through a digital connection to other communities and individuals; and being in the network through embeddedness in the web's global mediascape

(Androutsopoulos, 2013). Networked bilingualism entails user involving many linguistic resources while adhering to the restraints of written language in digital technologies, network resource accessibility, and orientation to audiences in the same network. Androutsopoulos (2013) has proven that a combination of CMC and DCOE can be as real as any physical setting and comparable to conventional ethnography research. Hence, this paper aims to emphasise that DCOE in various CMC settings can identify expressions and relationships between speakers.

Manglish

This section discusses aspects related to Malaysian English (ME) as background to the researcher's (first author) study. The type of online data discussed here is a non-native variety of English used by Malaysians, namely, Manglish. Manglish belongs in the basilectal group, a form of colloquial Malaysian English which differs in terms of phonology, syntactic, and lexical aspects. This non-native English is influenced by local languages to give users a sense of locality (Baskaran, 2005).

The linguistic features of Manglish include the use of particles, innovative spellings, code-switching, code-mixing, and borrowed words. The researcher's WhatsApp data consists of various Manglish features, mainly particles such as *lah*, *lor*, *ler*, and *meh* and others. Similar to other Manglish studies, the data also includes spelling modifications, such as 'gud' (good), 'ma' (my), 'bufday' (birthday), and 'setel' (settle) (Stapa & Shaari, 2013). Hassan and Hashim (2009) denote that the use of loan words is another feature of Manglish. For example, Manglish often borrows words from local languages, including Malay, Chinese, Tamil, and indigenous languages, such as 'kampung', 'matcha', 'angpow', and 'ngabang', since there are no English equivalent words that represent these intended meanings (Hassan & Hashim, 2009). In general, the features of Manglish are observable at various linguistic levels in both face-to-face and online communication with various positive and negative functions (Tay et al., 2016). However, the researcher's study looks beyond these functions and argues that there is more to them: the emergent aspect of Manglish's social meaning.

In Malaysia, there have been several studies conducted within various CMC settings (Stapa & Shaari; 2013; Tay et al.; 2016; Rusli et al., 2018). However, most of these studies collected Manglish features found online and listed its lexical functions. Although these researchers put forward the idea of a relationship between Manglish and youth identity, their observations were limited to textual data. This paper argues that understanding Manglish and its social meaning requires more than looking at the functions of conversational features. Manglish should be explored in correspondence to what the features actually signal in conversations and to the interlocutors. The significance of adequate contextualisation should be emphasised for building an understanding of Manglish's use in socio-cultural and digital contexts.

Methodology

The remainder of this paper is a reflection of online data collection methods. It provides an overview of participants and data collection methods, presenting the important steps of online ethnography. The analysis of extracts of conversations, focusing on information obtained by the researcher are outlined. Throughout, the aim is to demonstrate how the

parameters of online ethnographies are manifested in the participants' online WhatsApp chats.

Participants and Data Collection

The researcher collected data between July 2015 and January 2017 with a total of 714,999 words. Data consisted of 248 sets of naturally occurring WhatsApp conversations obtained from one-on-one existing WhatsApp chats to prevent pre-planned interactions. In total, 52 participants between the ages of 18 and 25 participated in the researcher's study. The participants come from various ethnic backgrounds, representing the three major ethnic groups: Malay, Chinese, and Indian. Two examples are used to demonstrate speech style variety and social meaning.

The researcher's study involves different stages of online ethnography, including establishing relationships with the participants (Kytölä & Androutsopoulos, 2012), systematic observation of online activities, and contact with participants (Androutsopoulos, 2008). The following section highlights the importance of establishing relationships with participants when conducting an online ethnographic approach. Kytölä and Androutsopoulos (2012) in their *Futisforums* study emphasise the importance of the researcher being 'initially involved and getting acquainted with the community as a casual visitor' (p. 184). It is believed that this is important to ensure access to the participants' life histories and gain reliable data interpretation.

Positioning Researcher's Youth Identity to Create Online Rapport

Mann (2021) discusses the importance of understanding one's role as a researcher in acquiring reliable data. Researcher positionality, such as age, gender, race, and class, affects the relationship between the researcher and participants, thus affecting the nature of the data.

During this data collection process, the researcher analysed the identity construction and Manglish language style of Malaysian youths in WhatsApp chats. The researcher shows how features such as *lah*, *lor*, *leh*, and *de* connote specific social meanings, such as intimacy and ethnicities. It is interesting that she sought to collect not only naturally occurring conversations, but also the metapragmatic and social output of these conversations. To do so, the researcher invested time into developing familiar relationships with the participants. It is argued that she is actually aligning her youth identity with the participants. Such a practice is demonstrated in Mann's (2021) study on Javanese youth identity, in which he recruited young Javanese research assistants who could transcend the role of researcher and connect with the participants, as these assistants were similar in age to the study subjects. Mann developed intimacy with the young participants by creating a sense of belongingness through linguistic choices and small talk. Similarly, the researcher engaged in a series of informal talks with the participants and disclosed personal information about herself, such as her studies, travel experiences, and life as a student in Liverpool. This created intimacy, as her age and youth identity resonated with the participants. We believe that such personal sharing is an important step prior to data collection, as it creates common ground with the participants without shaping or redirecting specific responses.

Given (2008) names dialect as another attribute that determines the relationship between researchers and participants. In the WhatsApp chats, this researcher employed Manglish features, such as sentence final particles and informal online languages containing innovative spelling (such as 'perrrrfect') with excessive use of emoticons, to build rapport and elicit responses to particular linguistic/online social behaviours. These features were intended to make the process of eliciting information more casual and less tense for the participants. It is suggested that familiarization with youth language, such as emoticons and Internet language, is important in developing rapport and avoiding misunderstanding with the young participants during digital data collection (Gibson, 2020; Jowett et al., 2011). Similar researcher positionality is observed in the methodology of Jowett et al.'s (2011) study of sexuality and health, which emphasizes the alignment of participant and researcher identity in digital data collection. The researchers' sexual identity gave them an 'insider' status which orientated with the participants' identity. In the same line, the researcher's engagements and involvement with the participants grafted in-group solidarity: she is Malaysian, and she was in her 20s during the data collection process, creating a small age gap with the participants who ranged from 18 to 25. These similarities likely influenced her rapport with the participants. However, Jowett et al (2011) noted the limitation of their rapport with participants, since they only engaged with them during the interview process. This was not an issue for the researcher's study, as she and the participants communicated each other over an extended period of time. This paper claims that such commitment makes the development of rapport in online research more pertinent, ensuring the success of online ethnographies.

Aligning with the participants' youth identities also requires a degree of involvement or participation in the speakers' social activities. Third wave variationist studies have often witnessed the involvement of researchers being physically present in social settings to connect with participants (Lampropoulou, 2012). This paper proposes that these social activities are not limited to offline settings but can be achieved digitally as well. In her study, the researcher helped some of the participants in their daily lives. For example, one participant sought her help in answering linguistic tasks and other participants consulted her in choosing the best signature. At some points during the data collection, the researcher shared the linguistic patterns that she identified in conversations.

Overall, the above tactics established online rapport with the participants. As a result, the researcher obtained full cooperation from some of the participants regarding information on context, the background of their interlocutors, and other parties discussed in the conversations. For example, a female participant, Laura, openly shared her views and social distance from her interlocutors, which contributed to an ethnographic understanding of the roles and practices between the participants. This openness supports the use of online ethnographies when dealing with young participants.

The next section discusses WhatsApp observations in relation to the researcher's experiences in conducting online observations and how they contribute to successful ethnographic findings.

Discourse Centred Online Ethnography (DCOE)

In this subsection, DCOE approaches that the researcher adopted and adapted in her Manglish study were outlined. These approaches aim to pull all the details together to make sense of the context and findings.

i. Systematic Observation

The researcher's study adapted Androutsopoulos' (2008) approach to online ethnography by observing the speakers' activities, or what they 'do', in WhatsApp. This is known as 'systematic observation', which, in the context of her study, means that the entire sent conversation (previous to present) was read. Such an approach allows her to observe the frequency of the chats, the topics of interaction, and the use/effect of Manglish features. The researcher shifted her role to that of a silent reader, asynchronously observing the conversations. It is believed that such observations should be done after the conversations take place, since the researcher's participation in the chats in real time affects the authenticity and openness of the conversations. When observing the chats, she noticed that the participants shared personal information, jokes, photos, and videos, as well as their emotions. WhatsApp chats, similar to other messaging platforms, creates a sense of control and intimacy among young speakers. They share personal information, enhancing the in-depth discourse analysis of the qualitative data.

ii. Maintain Openness

Androutsopoulos (2008) also suggests that the data should be viewed several times. Similarly, the researcher (2020) regularly browsed through the whole chats, not limited to specific conversational episodes. Such a step allows her to identify the core participants in her study. For example, she identified how specific speakers' use of Manglish features changed in several contexts. The use of Manglish increased or decreased when interacting with friends of different intimacy levels. The researcher also identified topics of interest that coincided with Manglish practices among the speakers, such romance and relationships. She also noticed a difference in frequency of contact between the speakers. This is a significant observation, as it coincides with social distance, which can lead to solidarity.

iii. Contact with Internet Actors

In Androutsopoulos' (2008) DCOE approach, after reading and observing the whole conversations, contact should be made with the Internet actors -- in this case, the participants. He emphasises initial contact, which, in the researcher's case, was done earlier via WhatsApp chats with regards to ethical issues and establish a rapport. Although there is variation in Androutsopoulos' interview methods (online and face-to-face), it is argued that interviews can also be fully conducted via online settings. The researcher demonstrates many benefits of online interviews. For instance, she was able to consistently contact the participants throughout her studies to clarify information from the chats. This included asking specific questions about why and when they used features such as *lah*, *lor*, *leh*, and *de* in their conversations. Furthermore, physical presence is not required from either party, creating a relaxed atmosphere and avoiding any possible embarrassment that might exist in a face-to-face context (Barton & Lee, 2013). The researcher's study proved fruitful, as she provided participants with screenshots of the conversations for further clarification and elicit awareness of their chosen linguistic styles.

Findings and Discussion

This section presents several advantages of the online ethnographic approach. As discussed in the previous section, participating in the researched community requires focus, commitment, and consistency. A successful online ethnographic approach enables insights that could not be gained without physically being there. Presenting two extracts of conversations from the researcher's study, the discussion illustrates the advantages of online ethnography. The examples mainly involve uses of Manglish in WhatsApp chats.

The online ethnographic approach conducted in the researcher's study led to a significant finding, as demonstrated in Table 1 below. During the systematic observation phase, the researcher found consistency in the uses of *lor*, *leh*, and *de* among speakers of similar ethnicities. Specifically, she detected excessive usage of *lor*, *leh*, and *de* when Chinese speakers interacted with those of similar ethnicities. However, further observation of the chats revealed that these features are also influenced by the social distance between the speakers. The researcher traced different representations of *lor*, *leh*, and *de*, even when the participants were interacting with speakers of similar ethnic groups. This shows that the aforementioned features are not limited to ethnic-attributed features, but also indicate other characteristics, such as in-group solidarity. Key to this example is when the researcher mapped the conversations of one participant, Laura, with interlocutors Xora, Jezmine, and Aisya who are studying Environmental Sciences at the same university in Malaysia. Table 1 presents the distribution of *lor*, *leh*, and *de* in Laura's conversations.

Table 1

Distribution of lor, leh and de per 1,000 words in Laura's Conversations (adopted from the researcher, 2020)

Ethnicity	Speaker-Addressee	Relationship with speaker (Laura)	Lor	Leh	De	Total no of features	No of words
			n/1000 words	n/1000 words	n/1000 words		
Chinese-Chinese	Laura and Xora	Course mates and friends	5	4	9	10	560
Chinese-Chinese	Laura and Jezmin	Course mates	5	9	3	75	4303
Chinese-Malay	Laura and Aisya	Colleagues	0	0	0	0	1251

The table shows that usage of *lor*, *leh*, and *de* occurs in Laura's conversations with Xora and Jezmine. Although Xora and Jezmine are Chinese speakers, the researcher traced different social functions of *lor*, *leh*, and *de* in Laura's conversations with both speakers. For instance, Laura and Jezmin demonstrated ethnic identity through the use of *lor*, *leh*, and *de* when interacting with each other.

In 2016, Laura, who had expressed concern with group assignments, initiated a discussion with Jezmine about tasks and reports related to entrepreneurship courses. The conversation began with an inquiry as to whether Laura knew how to contact their other course mates about the group assignment, all of whom appeared to be Malay. In this conversation, the researcher traced several indicators of in-group and out-group markers, which she clarified with the speakers.

Jezmine: Cuz we dun even can contact them...

Because we can't even contact them

Laura : dun know **leh**....the others all malay 😊

*I don't know....the others are all **Malay***

Jezmine: **Yalor**, very annoying **le**

Yes, it's very annoying

The first indicator is Jezmine's categorisation of 'everyone' as *we* and *them* in the conversation. The pronoun creates distance between Jezmine and Laura from other ethnic groups (Ciotti, 2016). Moreover, the pause in Laura's sentence represents a perception towards Malays, which is further reinforced through the 'Neutral Face' emoticon that implied negativity toward the ethnic group. Jezmine aligned with Laura's negative emotions by agreeing with the statement, 'Yalor', and supporting it with negative adjectives, 'very annoying le'. The researcher suggested that *leh* in this context does not represent a racialized slur, but rather a way of 'othering' or pejorative implications by using varieties of Manglish. The openness from both Laura and Jezmine in distancing themselves from the Malay ethnic group is, therefore, triggered by their ethnic identity.

In Laura and Xora's conversation, intimacy is represented through these features instead of ethnic identity. The researcher was able to identify their close relationship through their student life and activities on campus. Laura also openly revealed her closeness with Jezmine and Xora to the researcher, as Laura regards has a familiar relationship with the researcher. The conversation illustrates Xora's plans to bathe at Laura's accommodation following their elective class, as the water supply in Xora's hostel block is disrupted. The conversation diverges to other topics, such as skipping classes.

Xora: Why skip class o

Why skip class?

Laura: i go out **leh**, haha

I want to go out leh, haha

but u can come also **lor**

but you can come also lor

just dun go block d 😞

just don't go to block D

Xora: Haha, ok lor 😊

Haha, okay lor

In this conversation, Laura and Xora's intimacy is represented through their vocativeness and self-disclosure. Laura openly announces that she is going to skip class and invites Xora to join her. Xora's insistence to come to Laura's hostel despite knowing that Laura will be absent further clarifies their intimacy: Xora not only regards Laura as a course mate, but also as her friend. Laura engages with Xora's difficulties, as she proposes other alternatives to her hostel room. That is, Xora can follow her roommate back to the hostel. 'Lor' in this context acts as an emphatic symbol, a mitigation strategy to Laura's prior excuse (that she would not be in the hostel). This reduces the social distance between the speakers. The extended conversations also reveal Xora's concern for and advice about Laura skipping classes. Their conversations are intertwined with jokes and laughter, which reinforcing the closeness of relationship.

From these examples, it is suggested that the use of *lor*, *leh*, and *de* were deliberate and that its major functions are as in-group and ethnic markers. Laura employs *lor*, *leh*, and *de* when interacting with Jezmine to portray her 'Chinese-ness' despite their social distance, while she uses the features to enhance/maintain intimacy with Xora. The researcher's data suggests no occurrence of Chinese specific features (such as *lor*, *leh*, and *de*) when Laura is communicating with her Malay friend, Aisya. This reflects the racial gap between them.

It is unsurprising that ethnic origin can affect the feature usage among speakers. In this respect, this study's findings correspond with other sociolinguistic studies on dialect use between in-group and out-group members (see among other Bin Tahir et al., 2020; Paladino and Mazurega, 2020; Tararova, 2019; Hernandez, 2009). For example, Hernandez (2009) compares the use of nasal velarization between interviewers from different ethnic groups: Mexican Spanish and Salvadoran Spanish. Salvadoran participants interviewed by a Salvadoran-Spanish interviewer tended to employ a higher rate of nasal velarization, as compared to when the same informants interacted with a Mexican interviewer. Therefore, linguistic features play a significant role in identifying the members of an ethnic group, and they are often negotiated (reduced) to accommodate the outgroup speakers.

This paper intends to highlight how the researcher came to such conclusions through an online ethnographic approach. Despite knowing the researcher's Malaysian ethnicity, the rapport established earlier allowed Laura to be open and comfortable in sharing her true feelings regarding the racial gap shown in the conversation. Laura revealed this information after the researcher familiarized herself with the chats. These are examples of the importance of creating an online presence and connection with the respondents to procure genuine responses. Therefore, this paper has successfully demonstrated how an ethnography in a digital setting can inform a textual interpretation of the data.

Conclusions

In conclusion, this paper offers both empirical and methodological contributions to the third wave of variationist studies. Empirically, the researcher study finds that the uses of Manglish extend beyond the conventional functions of the features (e.g., to mitigate, to highlight). Instead, she illustrates the social meaning, or meso-level social categories, of the Manglish used in WhatsApp chats among youths. Specifically, she demonstrates how Manglish features such as *lor*, *leh*, and *de* can indicate ethnic identity or in-group solidarity among Chinese speakers. This finding aligns intercultural research showing that local language or variety plays a significant role as unifier of a multicultural society (see for instance Bin Tahir et al., 2020).

Following these findings, this paper expands on the researcher's approach to DCOE (Androutsopoulos, 2008), highlighting how she successfully positioned her youth identity to create rapport and gain trust with the participants. Throughout the process, the researcher maintains openness by not limiting her interest to certain topics but by examining all communication exchanges with the participants. The researcher was not aware of her approach, but it allowed her to interpret the participants' stances and nuances of the discourse's social meaning. The ethnographic approach also allowed the researcher to understand conversational background to interpret the participants' particular styles of speaking. In line with Androutsopoulos' (2008) framework, it is argued that the researcher's involvement provided an advantage to developing a grip of the linguistic patterns. This paper also highlights that a successful ethnographic approach involves systematic observation from the researcher and commitment from both the researcher and the participants for understanding tacit knowledge of underlying linguistic practices.

One of Androutsopoulos' (2008) ethnographic strategies is to make contact with Internet actors. He suggests that direct contact with participants allows for an in-depth understanding of the textual data. It is argued that the researcher's study aligns with Androutsopoulos' idea in that she needs to confront the participants and clarify conversations. The researcher confirmed the underlying meanings of the conversations, including emoticons employed by the participants. Narratives told online, such as Laura's, challenge the traditional ethnographic format and debates surrounding online ethnography. It seems that the new generation of Internet users, specifically youths, prefer openness when interacting in online settings. Hence, this paper proposes that these methodological aspects are directly applicable when investigating youths in private or multi-authored chats in both synchronous and asynchronous settings.

References

- Abidin, C., & De Seta, G. (2020). Private messages from the field: confessions on digital ethnography and its discomforts. *Journal of Digital Social Research*, 2(1): 1–19.
- Androutsopoulos, J. (2006). Introduction. Special issue on Sociolinguistics and computer-mediated communication. *J. Socioling.* 10 (4), 419–438.
<https://doi.org/10.1111/j.1467-9841.2006.00286.x>.
- Androutsopoulos, J. (2008). Potentials and Limitations of Discourse-Centred Online Ethnography. *Language@ Internet*, 5.
<https://www.languageatinternet.org/articles/2008/1610>
- Androutsopoulos, J., & Beißwenger, M. (2008). Introduction: Data and methods in computer-mediated discourse analysis. *Language@ internet*, 5(2).

- <https://www.languageatinternet.org/articles/2008/1609/introduction.pdf>
- Androutsopoulos, J. (2013). Online Data Collection. In C. Mallinson, B. Childs, & G. Van Herk (Eds.), *Data collection in sociolinguistics: Methods and applications* (pp. 236–250). New York: Routledge.
- Androutsopoulos, J. (2013). Networked multilingualism: Some language practices on Facebook and their implications. *International Journal of Bilingualism*, 19(2), 185–205. <https://doi.org/10.1177/1367006913489198>
- Bamman, D., Eisenstein, J., & Schnoebelen, T. (2014). Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, 18(2), 135-160. <https://doi.org/10.1111/josl.12080>.
- Barton, D., & Lee, C. (2013). *Language online: Investigating digital texts and practices*. London: Routledge.
- Baskaran, L. (2005). *A Malaysian English Primer: Aspects of Malaysian English Features*. Kuala Lumpur: University of Malaya Press.
- Bengtsson, S. (2014). Faraway, so close! Proximity and distance in ethnography online. *Media, Culture & Society*, 36(6), 862-877.
- Boccagni, P., & Schrooten, M. (2018). Participant observation in migration studies: An overview and some emerging issues. *Qualitative research in European migration studies*, 209-225.
- Boellstorff, T. (2015). *Coming of age in Second Life*. Princeton University Press.
- Bolander, B., & Locher, M. A. (2014). Doing sociolinguistic research on computer-mediated data: A review of four methodological issues. *Discourse, Context & Media*, 3, 14–26. <https://doi.org/10.1016/j.dcm.2013.10.004>
- De Seta, G. (2020). Three lies of digital ethnography. *Journal of Digital Social Research* 2(1): 77–97.
- Döring, N. (2003). *Sozialpsychologie des Internet*. 2nd ed. Göttingen: Hogrefe.
- Eckert, P. (2019). The limits of meaning: social indexicality, variation and the cline of interiority. *Language* 95(4). 751-76. <https://doi.org/10.1353/lan.2019.0072>
- Gibson, K. (2020). Bridging the digital divide: Reflections on using WhatsApp instant messenger interviews in youth research. *Qualitative research in psychology*, 1-21. <https://doi.org/10.1080/14780887.2020.1751902>
- Given, L. M. (2008). *The SAGE Encyclopaedia of qualitative research methods* (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. <https://doi.org/10.4135/9781412963909>
- Gobo, G. (2008). *Doing ethnography*. Los Angeles; London: Sage.
- Greschke, H. M. (2007). Bin ich drin? – Methodologische Reflektionen zur ethnografischen Forschung in einem plurilokalen, computervermittelten Feld. *Forum: Qualitative Social Research*, 8(3). <http://www.qualitative-research.net/fqs-texte/3-07/07-3-32-d.htm>
- Grills, S. (1998). An invitation to the field: Fieldwork and the pragmatists' Lesson. In S Grills (ed) *Doing Ethnographic Research: Fieldwork Settings*, Thousand Oaks, CA: Sage.
- Hassan, N., & Hashim, A. (2009). Electronic English in Malaysia: features and language in use. *English Today*, 25(04), 39. <http://doi.org/10.1017/S0266078409990435>
- Illingworth, N. (2001). The Internet matters: Exploring the use of the Internet as a research tool. *Sociological research online*, 6(2), 79-90. <https://doi.org/10.5153/sro.600>
- Jowett, A., Peel, E., & Shaw, R. (2011). Online interviewing in psychology: Reflections on the process. *Qualitative Research in Psychology*, 8(4), 354-369. <https://doi.org/10.1080/14780887.2010.500352>

- Kytola, S., & Androutsopoulos, J. (2012). Ethnographic perspectives on multilingual computer-mediated discourse: insights from Finnish football forums on the Web. In *Multilingualism, discourse, and ethnography* (pp. 181-198). Routledge.
- Lampropoulou, S. (2012). *Direct speech, self-presentation and communities of practice*. London: Continuum.
- Ling, R., & Baron, N. S. (2007). Text messaging and IM: Linguistic comparison of American college data. *Journal of language and social psychology*, 26(3), 291-298. <http://doi.org/10.1177/0261927X06303480>
- Moore, E. (2003). *Learning Style and Identity: A Sociolinguistic Analysis of a Bolton High School*. Unpublished PhD dissertation. Manchester, U.K.:University of Manchester.
- Moore, E., & Podesva, R. (2009). Style, Indexicality, and the Social Meaning of Tag Questions. *Language in Society*, 38(4), 447–485. <http://doi.org/10.1017/S0047404509990224>
- Rusli, M. F., Aziz, M. A., Aris, S. R. S., Jasri, N. A., & Maskat, R. (2018). Understanding Malaysian English (Manglish) Jargon in Social Media. *Journal of Fundamental and Applied Sciences*, 10(2), 116–125. <http://dx.doi.org/10.4314/jfas.v10i2s.10>
- Mohamad, S. N. H. (2020). *Social Meaning, Indexicality and Enregisterment of Manglish in Youth WhatsApp Chats* [Doctoral Dissertation, The University of Liverpool]. ProQuest Dissertations Publishing. <http://search.proquest.com.ezaccess.library.uitm.edu.my/dissertations-theses/social-meaning-indexicality-enregisterment/docview/2447591512/se-2?accountid=42518>
- Stapa, S. H., & Shaari, A. H. (2013). Language Transformation and Innovation in Online Malaysian. *Pertanika Journal of Social Sciences & Humanities*, 21, 131–144. Available from https://www.researchgate.net/publication/289581971_Language_transformation_and_innovation_in_online_Malaysian_English
- Tay, L. C., Chan, M. Y., Yap, N. T., & Wong, B. E. (2016). Discourse Particles in Malaysian English: What Do They Mean? *Bijdragen Tot de Taal-, Land- En Volkenkunde*, 172(4), 479–509. <http://doi.org/10.1163/22134379-17204002>
- Thi, L. S. (2016). Ethnography In The Cyberspace: Problems And Prospects. *International Journal of Information Systems and Engineering*, 4(2), 43–48. <https://doi.org/10.24924/ijise/2016.11/v4.iss2/43.48>
- Tsiplakou, S. (2009). Doing (bi) lingualism: Language alternation as performative construction of online identities. *Pragmatics*, 19(3), 361-391. <https://doi.org/10.1075/prag.19.3.04tsi>
- Underberg, N. M., & Zorn, E. (2013). *Digital ethnography: Anthropology, narrative, and new media* (1st ed.). Austin, TX: University of Texas Press.