Uncovering the Lexical Variation and Change in Malaysian English:
A Corpus-Driven Analysis

Noraishah Gulnazir [a]
noraishah.g@help.edu.my
HELP University, Malaysia

Khazriyati Salehuddin [b]
khazudin@ukm.edu.my
Center for Research in Language and Linguistics,
Universiti Kebangsaan Malaysia, Malaysia

ABSTRACT

Malaysian English has been known to be influenced by local traits and characteristics as well as global influences and trends. Despite being traditionally based on the British English, Malaysians today seem to show increasing prominence in the use of American English due to its widespread dissemination on media. Thus, to investigate the influence of American English on Malaysian English, this study examined the usage of two American lexical items that emerged in American Twitter in 2013, namely, *lit* and *on fleek*, in Malaysian English. By using Twitter API Tools to obtain data, this research investigated 19,050 tweets totalling 203,482 tokens with *lit* and *on fleek* from the year 2013 until 2021 in Malaysian Twitter. Data were analysed through frequency distribution and concordance analysis via AntConc version 3.5.9. Additional tools such as Google Trends and Urban Dictionary were also utilised. The results reveal that *lit* and *on fleek* increased in usage in Malaysian Twitter in 2016 and 2017 respectively, two to three years after these lexical items first emerged in American English. Results also indicate that Malaysian Twitter users adopted these lexical items with the same precision in meaning as American English and incorporated them into Malaysian English. Aspects of semantic change or semantic drift are also revealed; these lexical items have transformed in terms of meaning and usage from its original form. This study supports the World System of Englishes theory by proving that lexical items from American English, as a hyper-central language, could spread to other varieties of English including Malaysian English. The research also demonstrates the concept of language mobility on social media, whereby lexical items from American Twitter could traverse to Malaysian Twitter. This study highlights the extent to which Malaysian English is influenced by American English and illuminates the dynamics of English language variation and change.

Keywords: corpus-driven study; lexical variation; lexical change; Malaysian English; Malaysian Twitter

INTRODUCTION

English language, like any other languages, is constantly evolving. Time has revealed the ever-changing state of English, be it through the medium of its usage or spread, or through the many ways in which the changes take place. Language variation and change provide opportunities for researchers to further understand the multivalent ways of language use in the society. This is important as language is "a cultural product of a community of practitioners" (Brooks & Ragin, 2008, p. 514, cited in Chambers & Schilling, 2013) that does not take place in a vacuum; rather, it occurs in the context of collective activities.

---

[a] *Main author*
[b] *Corresponding author*

Traditionally, English language variation and change took place as a result of colonisation. Today, online spaces such as social media have led to language development and the shift of language practices and linguistic repertoire (Lantz-Andersson, 2018). Digital technology can motivate language change, support the study of language contact and change, or be a vehicle for change (Crystal, 2001). Consequently, numerous new word forms or neologisms have emerged from the social media and the Internet, and have now become ubiquitous, through various word formation processes such as metaphorization, generalisation, metonymisation, pejoration, and melioration (Traugott, 2017). For instance, the lexis *googling* which derives from the search engine Google, came into prominence in 2010, and has since then been generalised as a verb in the English language (Krazit, 2010). According to Crystal (2001), the constant emergence of new terms, such as *spam*, is driven by the urgent need to describe as-yet-unnamed events and experiences. Internet users are constantly looking for word forms to express their experiences, convey the essence of the digital world, and overcome the communication constraints of the medium.

Generally, when a new lexical item makes its way into the language, its overall frequency of usage rises over time following the S-shaped growth curve (Feltgen et al., 2017). Chambers and Schilling (2013) characterised the S-shaped curve of linguistic change into three separate parts, which are "initial stasis, rapid rise, and tailing off" (p. 32). Figure 1 shows the S-curve of linguistic change.
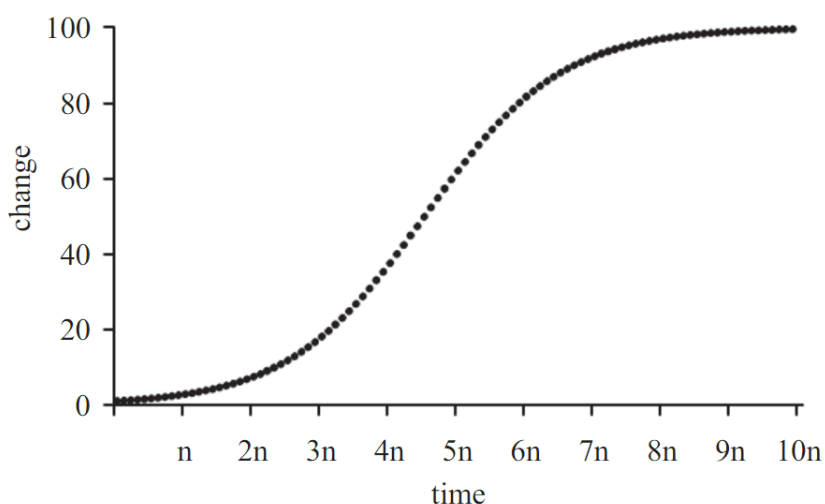

FIGURE 1. The S-curve of linguistic change in Tagliamonte and D'Arcy (2009)

Despite S-curve being the prototypical model in observing linguistic change, certain studies have not captured all three stages, and other studies have shown the opposite of the S-curve (Nurmi, 1999; Nevalainen, 2015). Additionally, emerging lexical items on social media have revealed that they either adhere to the prototypical S-curve or otherwise (Maybaum, 2013; Grieve et al., 2017).

Twitter (now known as X), a microblogging site that connects users through blogging and instant messages called *tweets*, is one form of social media that facilitates rapid and concise communication. As of the end of 2021, Twitter has 206 million users worldwide (Brian D, 2022). Because of its communicative nature, linguists have taken advantage of the wealth of linguistic data through the tweets; studies on language learning (Rosell-Aguilar, 2018), literacy practices (Gleason, 2018), sentiment analysis (Taboada, 2016), discourse analysis (Al-Ghamdi & Albawardi, 2020), and political communication (Santoso, Utari & Kartono, 2020) have been

conducted. Researchers (e.g., Huang et al., 2016; Grieve et al., 2017; Grieve et al., 2018; Grieve et al., 2019; Würschinger, 2021) have also used linguistic data from Twitter to explore language change and variation, particularly to understand the lexical items used in the tweets.

Twitter users are commonplace among Malaysians. Statistics show that there are 5.5 million Malaysian Twitter users in the year 2023 (Simon, 2023). Despite the fact that Malay is the national language of Malaysia and English is the official second language of the nation (Mahmud & Salehuddin, 2023), English, including the Malaysian variety of English (henceforth, Malaysian English) is widely used by Malaysians on social media, including Twitter (e.g., Rusli et al., 2018; Yunus, Zakaria, & Suliman, 2019). Like other varieties of English, Malaysian English has experienced language variation and change, initially as a result of colonialism, and more recently, due to the impact of globalisation (Gulnazir & Salehuddin, 2022). This variety of English is often described as a colloquial register that provides a sense of identity to its speakers and helps its speakers to build rapport among them. Malaysian English is characterised by phonological and semantic structures that are adopted from the various languages (particularly Malay) of Malaysia practiced by its multicultural and multiracial society (Salehuddin & Winskel, 2016), as well as aspects such as code-mixing, code-shifting, lexical shifts, localised cultural expressions, and simplification that reflects Standard British English (Thirusanku & Md Yunus, 2013; Pillai & Ong, 2018; Hashim, 2020; Salehuddin, 2022). Essentially, Malaysian English has emerged as a unique variety of English not only through the infusion of local traits and characteristics (Bolton, Botha & Kirkpatrick, 2020), but also through the impact of global influences and trends (Moody, 2020).

Although British English has always been the more dominant variety of English as a result of colonisation and its adoption into the Malaysian education system, American English today, like in other non-native countries, is gaining prominence in Malaysia among Malaysians (Shamsudin et al., 2019) due to the dominance of American popular culture in the media such as in movies, music, television (Gilquin, 2018) and now, digital communication (Crothers, 2021). Mair's theory of "The World System of Englishes" (2013) has labelled American English as a hyper-central variety of English, i.e., the hub for global English. British English, along with Australian English, Indian English, and Malaysian English, to name a few, has been labelled as a super-central variety of English due to the "massive presence of *lexical Americanisms*" in these varieties of English (Mair, 2013, p. 261). This results in globalisation of the language, where American English, with its influence in media and digital communication, acts as a unifying force across diverse English-speaking communities worldwide. Consequently, this form of globalisation leads to linguistic transformation (Tankosić & Dovchin, 2021), resulting in the formation of new words and expressions. Shifting the focus to social media, the vast number of online language users all around the word has contributed to rapid development in English language variation and change. This is because language change in social media is now driven by its global network of users, with the birth of global lexical items shared by global citizens.

PROBLEM STATEMENT

The usage of English language online in today's highly globalised world unfortunately, has also caused disparity. According to Martin (2020), instead of levelling the playing field for everybody, globalisation benefits a select few and neglects a large number of communities. Therefore, although as a local variety of English, Malaysian English should ideally be understood by all Malaysians, in reality, the Malaysian English that exists today is somewhat different from its forms years ago. Gen X (i.e., individuals born between 1965 and 1980), for example, may have difficulty in understanding the expressions used by Gen Z (i.e., those born between 1997 and 2012) which may have been caused by the influx of new global words into

the Malaysian English through the social media, including tweets. This echoes Lee's (2016) findings on the prevalent penetration of the English language due to globalisation, which has made the senior citizens in Korea feel linguistically inadequate. Such cases of language variation and change due to globalisation are in line with Pennycook (2007), which stated that "languages will flow and change around us, new combinations of languages and cultures will be put together, texts will be sampled and mixed in ever new juxtapositions" (p. 158).

Even though English is widely acknowledged to be prevalent in global computer-mediated communication, there is a dearth of research in studying the lexical units or grammar of English on social media (Coats, 2016). Narrowing it to the Malaysian context, generally, there is a paucity of research on lexical variation and change in Malaysian English. The recent studies have mainly focused on lexical features in a movie (Nor & Zamri, 2015), lexical borrowing in newspapers (Kunalan, Mutty & Francis, 2021), and slangs and jargons (Rusli et al., 2018; Izazi & Tengku Sapora, 2020). There is a need to investigate the presence of lexical items which are widespread online at a global scale in different varieties of English, for instance, in the Malaysian English, especially since the role of the English language in Malaysia has been driven by globalisation for a while (Azman, 2006). Hence, this study was conducted to explore the lexical variation and change in the Malaysian English, particularly on Twitter, by investigating two American lexical items that emerged in the American twitter in 2013, namely *lit* and *on fleek.* It aims at understanding how the global widespread and pervasiveness of American English influence Malaysian English on Twitter, especially in terms of lexical units, through the following research questions:

1. What is the trend of usage of both *lit* and *on fleek* in Malaysian Twitter from when it first emerged in American English in the year 2013 until the year 2021?
2. How are *lit* and *on fleek* utilised in Malaysian Twitter in terms of meaning and context of usage from the year 2013 until 2021? Do they show any similarities or differences with American English?

METHODOLOGY

This study aims to investigate two American lexical items - *lit* and *on fleek* - and their progression in Malaysian Twitter in terms of frequency and usage, from the year 2013 until 2021. These two lexical units were selected based on an earlier study by Grieve et al. (2017) which revealed that these lexical items emerged in American Twitter from the year 2013 onwards and have since then been circulating worldwide. To achieve this aim, the method proposed in Gulnazir and Salehuddin (2022) was adopted.

In corpus linguistic studies, the biggest distinction lies with the approaches taken, namely corpus-based or corpus-driven. Generally, corpus-based approach is a top-down approach that makes use of corpus data to challenge, support, or disprove theories. On the other hand, corpus-driven approach is a bottom-up approach which makes use of the corpus data to serve as the evidence to a research question. In corpus-based approach, there are preconceptions while in the corpus-driven approach there are no preconceived notions in investigating the data (Tognini-Bonelli, 2001). In corpus-driven approach, "the pure theory is hidden within the data" (Sinclair, 2004 cited in McEnery & Hardie, 2012, p.147) and the analysis of corpus is conducted by observing the keywords in context (KWIC) via concordance analysis.

The approach employed in this study is the corpus-driven approach. Therefore, to adhere to the corpus-driven approach, raw data of tweets with the two lexical items studied (*lit* and *on fleek*) were gathered from Malaysian Twitter from the year 2013 until 2021, and the data were then analysed through corpus methodologies. The lexical items were analysed in

terms of their frequency distribution, followed by the examination of usage of *lit* and *on fleek* through concordance analysis.

Generally, *lit* is a word form which has been institutionalised in the dictionary, and carries the following meanings: (1) *the past tense and past participle tense of the verb 'light'*, (2) *to illuminate,* or (3) *to ignite*. As with most other lexical items, the meaning of *lit* has expanded over the years. According to the Routledge Modern Dictionary of Modern American Slang and Unconventional English (Figure 2) by Dalzell (2018), *lit* emerged in 1899 in American English to carry the meaning *drunk*. The next definition for the word *lit* that is provided in the same dictionary carries the meaning *exciting* or *thriving;* this meaning emerged in the year 2014.

lit *adjective*
1 drunk *US, 1899*
- Like I said, he was always pretty well lit back in New York. — Raymond Chandler, *The Long Goodbye*, p. 209, 1953
- [T]he euphemisms for "drunkenness," e.g., "high," and "lit," from ten or fifteen years ago, and the direction of those euphemisms toward omnipotence[.] — William and Jerrye Breedlove, *Swap Clubs*, p. 151, 1964
- Nothing is more beautiful than four lit stooges in a graveyard spieling on[.] — John Nichols, *The Sterile Cuckoo*, p. 155, 1965
- — Collin Baker et al., *College Undergraduate Slang Study Conducted at Brown University*, p. 151, 1968

2 exciting, thriving *US, 2014*
- All the pretty girls where the money at / And it's lit, lit, lit, lit! –Joe Moses, "Lit." — *joe-moses-lit-lyrics.html*, 2015

FIGURE 2. The definition of *lit* in the Routledge Modern Dictionary of Modern American Slang and Unconventional English

However, according to Urban Dictionary (*urbandictionary.com*), *lit* now has several definitions. Aside from denoting excitement, it could also mean *cool* or *awesome*. Additional colloquial meaning includes *something that is turned up*, *popping*, *fire,* and *dope*. Although the initial definition recorded in the dictionary is only limited to *drunkenness* (as shown in Figure 2), in reality, *lit* is also used to describe intoxication caused by drugs.

Grieve et al. (2017) have found that the lexical item *lit* emerged in American English, particularly the American Twitter between 2013 and 2014. While the exact emergence and usage of this lexical item has not been explored in literature, *lit* is largely disseminated through popular culture. *Lit* began to be used by a famous Houston-based rapper *Travis Scott* as his signature line in his songs, by saying "It's lit!". While he did not invent the lexical item, he has championed the word *lit* in his music (Webb, 2018). This lexis is so common in popular culture that Google created a marketing report titled 'It's Lit: A Guide to What Teens Think Is Cool' in 2017.

The next lexical item investigated is *on fleek*. Similar to *lit, on fleek* also emerged in American English between 2013 and 2014 (Grieve et al., 2017). *On fleek* is synonymous to *on point*. As reported in Grieve et al. (2017), the origin of *on fleek* is less clear; it may be a genuine word invention, or it may also be connected to the word *flick*. However, what is evident is the fact that the usage of this individual word, which occurs as part of multi-word sequence, began after a 2014 video went viral. In the video, Peaches Monroe said "eyebrows on fleek" to describe her well-groomed eyebrows, which later went viral. The origin of this lexical item is also reflected in the definition provided by the Routledge Modern Dictionary of Modern American Slang and Unconventional English, as shown in Figure 3.
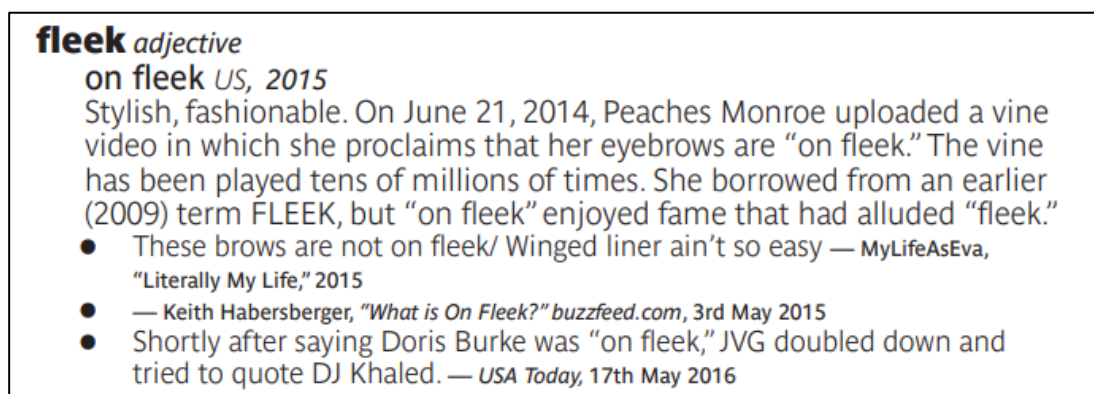
fleek *adjective*
on fleek US, 2015
Stylish, fashionable. On June 21, 2014, Peaches Monroe uploaded a vine video in which she proclaims that her eyebrows are "on fleek." The vine has been played tens of millions of times. She borrowed from an earlier (2009) term FLEEK, but "on fleek" enjoyed fame that had alluded "fleek."
- These brows are not on fleek/ Winged liner ain't so easy — MyLifeAsEva, "Literally My Life," 2015
- — Keith Habersberger, *"What is On Fleek?"* buzzfeed.com, 3rd May 2015
- Shortly after saying Doris Burke was "on fleek," JVG doubled down and tried to quote DJ Khaled. — *USA Today*, 17th May 2016

FIGURE 3. The definition of *on fleek* in the Routledge Modern Dictionary of Modern American Slang and Unconventional English

According to Grieve et al. (2017), the lexical item *on fleek* was strictly used to describe eyebrows when it first became viral. However, its usage has slowly expanded to cater to other forms of usages, for instance, to describe well-groomed hair.

DATA

To compile the data for this study (i.e., tweets with *lit* and *on fleek* in Malaysian Twitter) the Twitter API Developer Platform was utilised. With the ability to retrieve 10 million tweets per month, the API for Academic Research enables historical and real-time tweets to be mined. Due to its extensive, useful features, the Twitter Academic Access has been utilised in several linguistic studies, especially in investigating language variation and change (Grieve et al., 2017; Grieve et al., 2018; Grieve et al., 2019).

To search for tweets in English with *lit* in Malaysian Twitter, the following search criteria were used: (1) keywords: *lit* as the search term; (2) country: Malaysia; and (3) tweet language: English (only). The following query was generated:

"lit" place_country:MY lang:en.

The same was repeated with *on fleek*, using the following query:

"on fleek" place_country:MY lang:en.
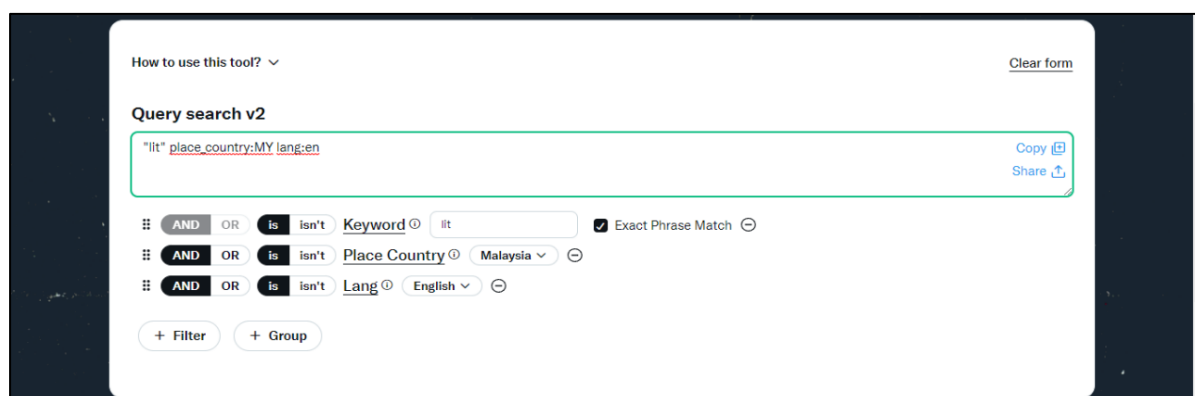
This is shown in Figure 4.

FIGURE 4. Using the Twitter Downloader Tool

Tweets were then downloaded in large batches for each year, i.e., from the year 2013 until 2021 in .csv format. Table 1 summarises the number of tweets found from the search.

TABLE 1. Number of tweets for both *lit* and *on fleek* in Malaysian Twitter

Lexical item	Number of tweets
Lit	18, 080
On Fleek	970
Total	19, 050

After compiling, the tweets were then cleaned manually from noise such as hashtags, links, and emojis. The following number of tokens were obtained from Tokenizer API after cleaning.

TABLE 2. Number of tokens for *lit* and *on fleek*

Lexical Item	Number of Tokens
Lit	192, 930
On Fleek	10,552
Total	203, 482

DATA ANALYSIS PROCEDURE

To answer Research Question 1, which is to investigate how *lit* and *on fleek* are distributed in the corpus and draw conclusions from them through frequency distribution, the frequency of usage of the lexical items *lit* and *on fleek* from the year 2013 until 2021 was first determined. This frequency was automatically derived from the Twitter API Tool, based on the year the tweets were tweeted. This type of frequency is known as either *raw* or *absolute frequency*, which refers to the number of times a lexical item occurs in the corpus.

Absolute frequency, however, is regarded insufficient (Baker, 2010) to determine the progression of the lexical items from the year 2013 until 2021. Therefore, in order to chart the distribution of the lexical items throughout the years, the number of times a lexical item occurs in the corpus has to be considered in relation to the total number of words i.e., tokens in the corpus. This is known as *relative frequency*. Therefore, to convert the raw or absolute frequency to the relative frequency, calculations according to the following formula need to be performed (Baker, 2010).

$$\frac{\text{Total number of occurrences of lexical item}}{\text{Total words in corpus (or sub-corpus)}} \times 1{,}000{,}000$$

FIGURE 5. Calculations for frequency per million words in Baker (2010)

The conversion of absolute frequency to relative frequency is called *normalisation*. Normalisation is necessary because it allows comparisons between corpora that have different sizes to be done accurately and fairly (McEnery & Hardie, 2011). The relative frequency is then used to determine the trend of frequency for both *lit* and *on fleek*, from the year 2013 until 2021, i.e., from the year these lexical items first appeared until the data was collected. This procedure helped determine whether these lexical items have been consistently used by Malaysians throughout the years. After the calculations for normalisation were performed, the data were visualised using a line graph, following Mat Awal et al. (2021).

Google Trends was also utilized to complement the results for the trends of usage on Twitter, following Grieve at al. (2017) and Jensen (2017). Since the data obtained above to trace the progression of *lit* and *on fleek* on Twitter were calculated manually, especially for the relative frequency through normalisation, it was helpful to have a database which generated similar data to identify the usage of *lit* and *on fleek* over the years from 2013 until 2021. Both synchronic and diachronic investigation of lexical items can be conducted on Google Trends; therefore, the earliest and latest occurrences of *lit* and *on fleek* in Malaysia were examined by utilising Google Trends.

To attain data for the present study, a number of criteria were customised on Google Trends. Firstly, both *lit* and *on fleek* served as the search terms. Next, the location was set to *Malaysia*. Since the data required was from the year 2013 until 2021, the time range selected was from *2013 until 2021*. Lastly, for the Category tab, *All* was selected for a comprehensive result.
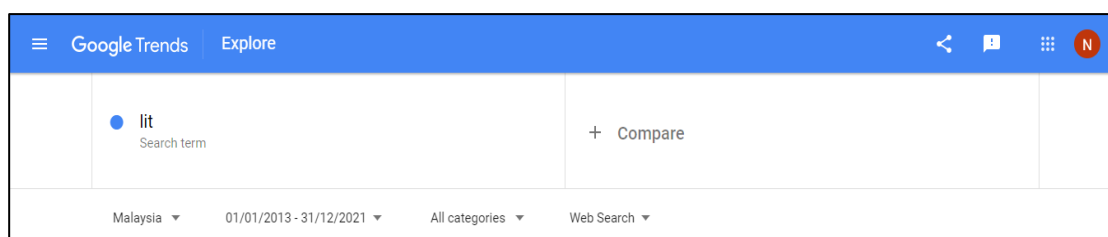

FIGURE 6. The criteria selected on Google Trends for the lexis *lit*

After the aforementioned criteria were determined, Google Trends automatically revealed the frequency of usage of these lexical items on Google during the specified time frame as well as the distribution and details of query (top and rising searches) associated to these lexical items in all the 13 states and 3 federal territories in Malaysia. All of these were visualised in geographical maps. By analysing the precise Google searches made for both *lit* and *on fleek*, this procedure provided support for the information gathered from Twitter and helped pinpoint the regions in Malaysia that may have aided in the spread of these new word forms (Jensen, 2017).

To answer Research Question 2, which concerns the usage and meaning of *lit* and *on fleek* in Malaysian Twitter particularly between 2013 and 2021, concordance analysis was performed in Antconc version 3.5.9. The present study followed Sinclair (2003) in terms of the analysis of the concordance lines. Sinclair (2003) outlined an extensive methodology in understanding the various meanings of a lexical item, which is summarised as follows.

When analysing concordance lines, the first procedure is to determine which word will be the *node* i.e., the target item, symbolised as N. For the lexical item *lit*, for example, the word preceding and succeeding the node were the co-texts, and they were denoted with either the (+) or the (-) symbol. One word situated to the right of the node word is denoted as N+1, while the first word to the left of node word is N-1. The node's recurrent co-occurrences with other words, also known as *co-texts*, have the tendency to develop a pattern. To ensure that the node word as well as the co-texts of *lit* and *on fleek* are given full attention to, in the Antconc software, the node word and the surrounding words on the left and right of the node word were highlighted and presented in different colours. This type of concordance analysis was conducted for both the lexical item *lit* and *on fleek* for each year from 2013 until 2021.

1	to give me answers for chinese lit
2	relationships, Love that keeps the flame lit.
3	so how...any idea to lit a candle?
4	merchant for 5mins, it starts pouring, lit a cigarette and put on my
5	Suyin with a super duper brightly lit cake hahaha!
6	get my results for the eng lit exam!
7	at night to see the brightly lit KLCC in front of our eyes.
8	iceberg contains more heat than a lit match.
9	First candle is lit on the First Sunday of Advent
10	Christmas Day in Singapore, brightly lit Orchard Road.
11	(1940) said Munsyi Abdullah was modern malay lit pioneer but Hassan Ahmad (1976) was with
12	back to doing lit review bye peasants.
13	We lit the diyas and offered prayers today.
14	bright the entire main road is lit up
15	I don't understand why people lit up cigarettes in the lift and
16	Finally got to lit up fire works!
17	The Eiffel Tower has been lit up in the colors of the
18	He lit up my life.
19	KL Towers lit up pink tonight for Breast Cancer
20	her eyes, about the way they lit up when she saw me.

FIGURE 7. Example of concordance analysis in the study

After identifying the different co-texts of the lexical items *lit* and *on fleek*, the co-texts were then grouped together to derive the possible meanings of the lexical items. Therefore, following the exact paradigm in Sinclair (2003), the co-texts were tabulated according to the number of concordance line as well as the category and the meaning of the co-texts, as illustrated in Table 3.

TABLE 3. Example of categorisation of meaning for *lit*

Line Number	Co-text	Category/Meaning
3, 9	Candle	Objects that illuminate light and
4, 15	Cigarette	ignite fire.
8	Match	
16	Fireworks	

After the meanings of the lexical items were interpreted based on the concordance lines, Urban Dictionary was used to verify and compare the usage of lexical items *lit* and *on fleek* in Malaysian English with the meanings in American English. The importance of Urban Dictionary in examining lexical items on Twitter is two-fold. Firstly, this user-generated site helps to track the earliest occurrence of words and phrases, and secondly, Urban Dictionary enables researchers to verify the definition of terms found in Twitter corpora (Grieve et al., 2017). Therefore, in the context of this study, Urban Dictionary was utilised to search for the definitions of the lexical items and determine whether Twitter users in Malaysia use both *lit* and *on fleek* similarly to American Twitter users, and whether they are referring to the same meanings when they do. Hence, similarity in the usage would suggest that Malaysian English is influenced by American English on Twitter. However, dissimilarity in the usage of both varieties of English would suggest that Malaysian English has not been influenced by American English with regard to these lexical items. To navigate through the Urban Dictionary page, only the search term, i.e., *lit* and *on fleek*, were required for the definitions to be revealed. This can be seen in Figure 8 below.
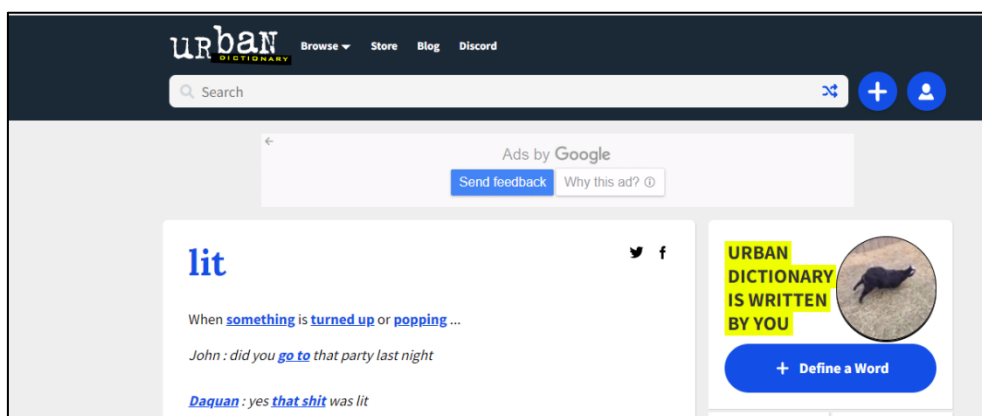
FIGURE 8. Navigating Urban Dictionary

RESULTS AND DISCUSSION

Research Question 1: *What is the trend of usage for both lit and on fleek in Malaysian Twitter from when it first emerged in American English in the year 2013 until the year 2021?*
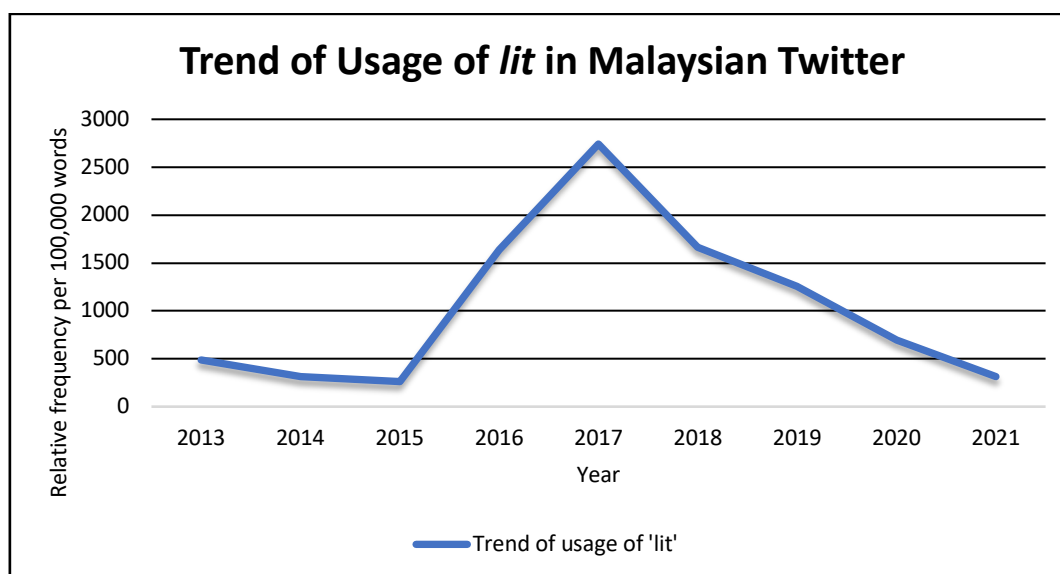

FIGURE 9. Trend of usage of *lit* in Malaysian Twitter from 2013 until 2021

The results for the trend of usage or frequency distribution of the novel lexical items reveal that for *lit*, the usage began with a slight decline between 2013 and 2015. In 2015, the lexical items expanded in its context of usage, and rose dramatically in 2016 before peaking in 2017. After the spike in 2017, the usage of *lit* dropped immediately and slowed down over the years until 2021. This finding is backed with the results from Google Trends, which also showed that the lexical item *lit* peaked in 2017.

The lexical item *lit*, as explained earlier can be divided into two separate groups of meanings. The first group is the conventional meaning, i.e., *the past and past participle of light* and the second group is comprised of the new meanings which have been invented in the American English, i.e., *being excited* etc. The expansion of the meaning of *lit* in Malaysian Twitter following the latter group takes place from the year 2015 onwards. The usage increased steadily until it peaked in 2017 before the immediate drop in 2018. After the abrupt decline, the usage of *lit* in Malaysian Twitter gradually diminished in frequency until 2021.
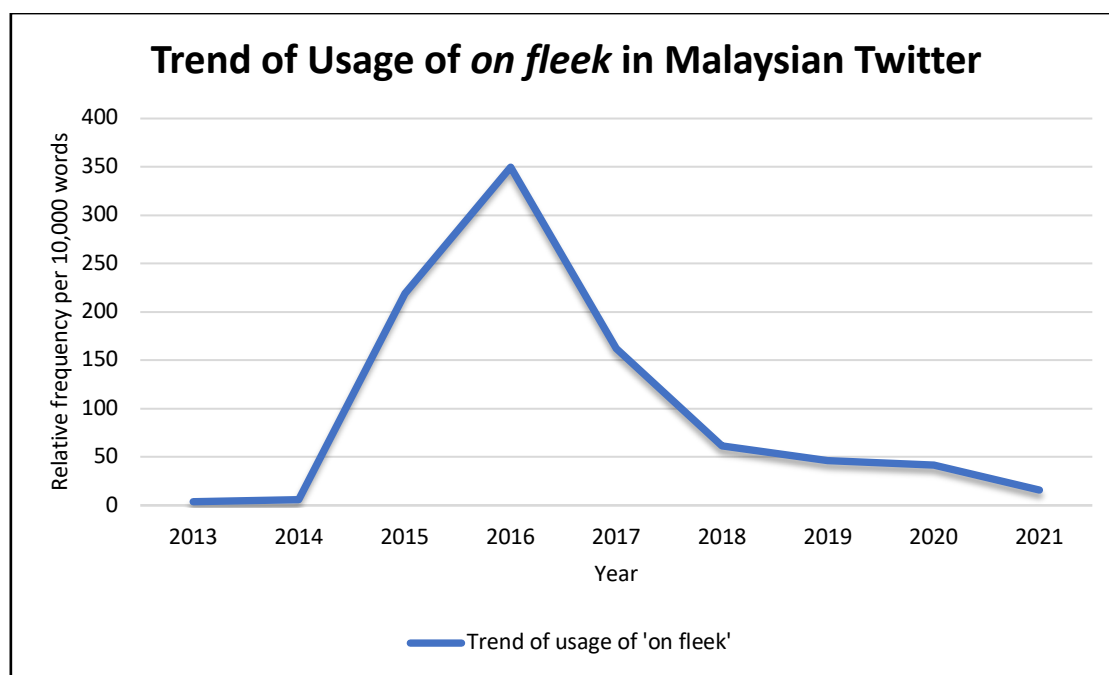
FIGURE 10. Trend of usage of *on fleek* in Malaysian Twitter from 2013 until 2021

The lexical item *on fleek* also shows a similar pattern, whereby the usage began minimally in 2013 and 2014, rose slowly, and finally peaked in 2016. The use of *on fleek* in Malaysian Twitter then decreased immediately, marking 2021 as the year with the lowest usage. This result is also complemented by data from Google Trends, which revealed that the usage of *on fleek* increased in 2016.

To summarise, the trend of usage of *lit* and *on fleek* in Malaysian Twitter between 2013 and 2021 is such that it begins with a slow progression but follows a linear pattern of growth, gradually increasing and peaking in 2017 (for *lit*) and 2016 (for *on fleek*) before declining. The conclusion which can be derived from Figure 9 and Figure 10 is manifold. First, it reveals that the novel lexical item *lit* which emerged in American English in 2013 took two years to be a part of Malaysian English and another year for it to become a prominent usage in Malaysian English. As for *on fleek*, the lexical item is used minimally in Malaysian Twitter between 2013 and 2014, i.e., the period when it first emerged in American English. Compared to *lit*, the speed of the spread of *on fleek* among Malaysians is slower, taking three years before peaking in 2016.

The present study is a diachronic study, with accounts for a total of nine years, from 2013 until 2021. Based on Figure 9 and Figure 10, for both the lexical items *lit a*nd *on fleek*, the immediate decline in the year 2018 and 2017 respectively indicates that the trend of frequency does not resemble the S-curve shape because the prototypical S-curve normally shows a gradual decline after a peak is reached. This is regarded as an *inverted S-curve* (Nevalainen, 2015) whereby the pattern shows reversal in its course immediately after the said course is picked up. Therefore, the curves of linguistic change for both *lit* and *on fleek* are inverted S-curves.

Additionally, the inverted S-curves of linguistic change revealed for both *lit* and *on fleek* suggest that the usage of these lexical items in Malaysian Twitter will gradually decrease after the year 2021. One possible reason is the shifting landscape of popular culture, as proposed by Pennycook (2007). Language often reflects the interests and trends of a society, and as popular culture changes, certain terms may lose relevance and usage. The English language is dynamic and is constantly evolving; therefore, it is common for lexical items such

as *lit* and *on fleek* to decrease in usages over time, especially on social media platforms like Twitter, as new terms emerge.

Apart from that, results from Research Question 1 also show that the highest query registered on Google Trends for both the lexical item *lit* and *on fleek* originates from Malaysians residing in the Federal Territory of Kuala Lumpur. This means that compared to the residents of other states of the country, Malaysian urbanites were the group who were most concerned with both *lit* and *on fleek*.

FIGURE 11. Distribution of *lit* and *on fleek* based on the regions in Malaysia

Kuala Lumpur or colloquially referred to as KL, is the capital city of the country and is regarded as an urban, metropolitan area. The city is also known as the hub of development in Malaysia. As elaborated earlier, Malaysian urbanites adopt and adapt readily in this era of globalisation given the preconceived rules and governance that have already favoured the urban population. To add to that, the rise of globalisation has also ensured that the urban population attains its benefits with the easy access to technological advancements and the Internet. Thus, with both the lexical items *lit* and *on fleek* being closely associated with urbanites in Kuala Lumpur, it could be concluded that with regard to the lexical items investigated in this study, only the urbanites, or at least those who are socioeconomically privileged and have the means to use social media, experience the effects of globalisation, particularly on Twitter. While *lit* reached states from West Malaysia such as Kuala Lumpur and Penang, as well as East Malaysia (Sabah and Sarawak), it recorded the highest usage in Kuala Lumpur, and *on fleek* succeeded to exclusively capture the attention and reach of only those residing in Kuala Lumpur. Therefore, this proves that with regard to the lexical items *lit* and *on fleek*, the effects of globalisation on Twitter are fully experienced only by urbanites in Malaysia. The findings of this study, which show that urbanites are most affected by language globalisation online, can be linked to the historical context of globalisation and the English language in the nation as posited by Azman (2006). The English language in Malaysia was initially intended for colonial governance; it later transformed into aspirations for nation-building, and currently for global competition. These developments, which have brought numerous benefits to the urbanites, have unfortunately left some in the rural communities behind.

Research Question 2: *How are lit and on fleek utilised in Malaysian Twitter in terms of meaning and context of usage from the year 2013 until 2021? Do they show any similarities or differences with American English?*
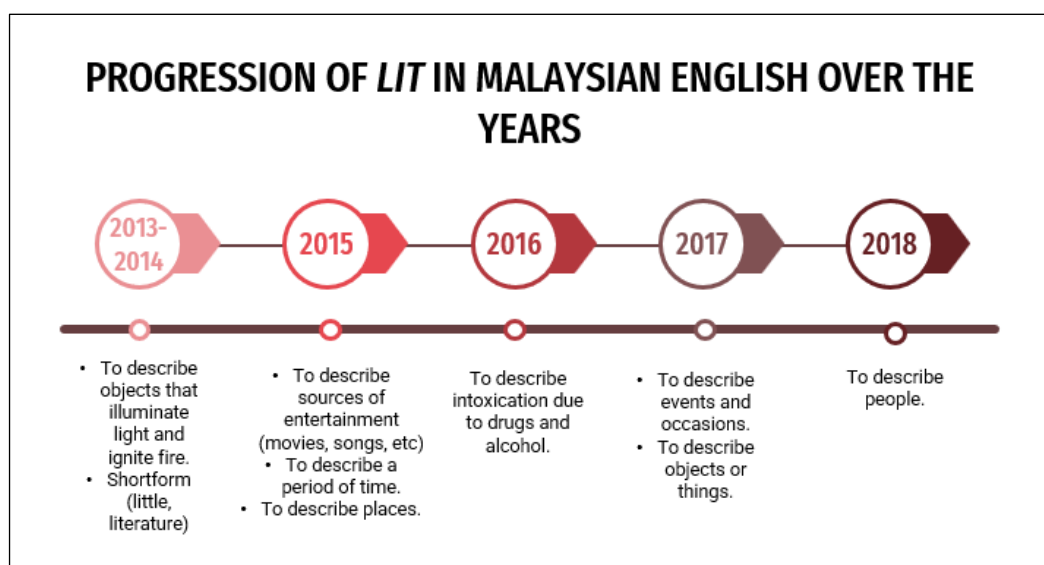

FIGURE 12. Progression of the meaning of *lit* in Malaysian Twitter over the years

Figure 12 shows the progression in terms of the meaning of the lexical item *lit* in Malaysian Twitter over the years. Between 2013 and 2014, the lexical item was mostly used within its original context, i.e., as the past and past participle tense of *light* and as a shortened form of the word *little*. From 2015 until 2021, the lexical item began taking new forms of meaning in Malaysian Twitter. *Lit* was then largely used to describe sources of entertainments, such as songs, films, dramas, and concerts. Malaysian Twitter users also used *lit* in relation to how exciting a certain period of time is, for instance, through the expression how *lit* last night was. Additionally, *lit* was used to describe the feeling of intoxication, caused by either drugs or alcohol. Events and occasions such as party and stage performances were also regarded as *lit*. The lexical item expanded even more by extending it to describe objects or things such as food or even smartphones which has exceptional qualities. Lastly, Malaysian Twitter users were found to associate *lit* with people, as a form of positive description.

FIGURE 13. Progression of the meaning of *on fleek* in Malaysian Twitter over the years

On fleek has been strictly used according to its meaning which is to describe something that *looks good and on point*. The only difference is that the usage first began by referring *on fleek* to well-groomed eyebrows, before the usage expanded to other contexts, for example, to refer to *cosmetics* and *hair* over the years.

The usages discovered from Malaysian Twitter were then compared to the usages available in American English on Urban Dictionary; it is found that the usages are identical with one another. This means that generally, Malaysians use the lexical item *lit* and *on fleek* in the same manner as American English.

As shared in The Routledge Dictionary of Modern American Slang and Unconventional English (Dalzell, 2018), *lit* is first defined in 1989 as *drunk*. It was only in 2014 that another meaning was registered for *lit* in the dictionary, which is *exciting*. Therefore, based on Traugott (2017), *lit* has undergone generalisation. The initial meaning which is *drunk* is now extended to other forms of intoxication, including drugs. *Lit* denotes excitement, but it has expanded in its context of usage. The lexical item is now used to refer to various events, occasions, sources of entertainment, and is also used to describe a period of time. The same applies to *on fleek*. As elaborated in Grieve et al. (2017), *on fleek* has become generalised over time. The usage was initially only restricted to describe the eyebrows, as recorded in Dalzell (2018), but the usage expanded. *On fleek* is now used to refer to "various other well-presented things, including makeup, braids and people" (Grieve et al., 2017, p. 120). This is similar to the usage in Malaysian Twitter, in which the users first associated *on fleek* with *eyebrows* but eventually included other nouns such as *hair*.

As shown previously, both *lit* and *on fleek* were used and disseminated widely on social media, particularly Twitter (Grieve et al., 2017). It is evident that Malaysian Twitter users utilised *lit* and *on fleek* in the Malaysian English in the same manner as users in the American English did. Thus, this study has proven that *lit* has undergone semantic drift to the extent that it has influenced other varieties of English. In this case, the American lexical items *lit* and *on fleek* have influenced and have brought forth the usage of these lexical items in Malaysian English. *Lit* is expanded from its original form as the past and past participle tense of *light* to a multitude of meanings, following American English, while *on fleek* is acquired as a novel lexical item which eventually became generalised.

The findings in this study can explain Mair's (2013) theory of The World System of English. According to Mair's model, American English has established itself as a hyper-central variety, having the greatest potential to impact other varieties of English. The key indicator of the influence of American English is when lexical items from American English spread into other varieties of English. This study has shown that Malaysian English, as a super-central variety, is indeed impacted by American English, with regard to both *lit* and *on fleek*. *Lit* and *on fleek* have traversed from American Twitter to Malaysian Twitter, proving that lexical variation and change is driven by digital media and globalisation (Martin, 2020; Tseng & Hinrichs, 2020).

CONCLUSION

This study has investigated the lexical variation and change in Malaysian English to understand the extent to which Malaysian English is influenced by American English. By adopting a corpus-driven approach using frequency distribution and concordance analysis, the results of this study have provided a general overview on the emergence of new lexical items such as *lit* and *on fleek* in Malaysian English and enabled the examination of the usage of these words online.

While the results obtained provide corroboration to previous studies, there are several limitations of the present study. It is necessary to highlight that the results for this study cannot

be generalised for other lexical items in Malaysian English. The pervasiveness displayed by American English in permeating Malaysian English is only limited to the two lexical items investigated in this study, which are *lit* and *on fleek.* Thus, future studies could perhaps investigate more lexical items from American English and examine its impact towards Malaysian English.

This study has shown that in this era of digital communication, it does not take long for lexical items from a hyper-central variety (American English) to penetrate a super-central variety of English (Malaysian English) in terms of their forms (i.e., spelling and pronunciation) and functions (i.e., meaning, and the context of usage). Interestingly, the high occurrence of these new lexical items seems to be only temporary; they have the potential to cease to exist in a few years when newer lexical items become more appealing to the speech community. This proves that digital media and globalisation plays an important role in driving lexical variation and change. Social media platforms such as Twitter, where English is the primary means of communication, are capable in contributing to the emergence and disappearance of lexical items.

The advancement of technology in driving digital communication is expected to result in the emergence of new words and phrases in the future. New applications and platforms will continually reshape the way language users interact with one another online, bringing new lexical items to the forefront. These novel lexical items would not be uniform in either their emergence or spread− they will likely be catered to different language user groups. Youths who are known to be well-acquainted with latest digital trends easily adopt new lexical items due to their online experiences. Despite this being the case for now, it is highly probable that language use online may affect the broader English-speaking community in the future. English language, with its dynamic nature, has always evolved with regard to technological changes and will continue to do so. Lexical variation and change remain a promising field of study, and it is hoped that this research has contributed valuable insights to this area of investigation.

REFERENCES

Al-Ghamdi, N. A., & Albawardi, A. H. (2020). Multivocality of Saudi COVID-19 discourse in social media posts: A socio-semiotic multimodal perspective. *GEMA Online Journal of Language* Studies, *20*(4), 228-250.

Androutsopoulos, J. (2011). Language change and digital media: a review of conceptions and evidence. *Standard Languages and Language Standards in a Changing Europe 1,* 145-159.

Androutsopoulos, J. (2014). Moments of sharing: Entextualization and linguistic repertoires in social networking. *Journal of Pragmatics*, *73*, 4-18.

Azman, H. (2006). English language in rural Malaysia: Situating global literacies in local practices. *3L, Language, Linguistics, Literature, 11*, 99-119.

Baker, P. (2010). *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Baskaran, L. M. (2005). *A Malaysian English Primer: Aspects of Malaysian English Features*. Kuala Lumpur: University of Malaya Press.

Baskaran, L. M. (2008). Malaysian English: Morphology and syntax. In R. Mesthrie (Ed.), *Varieties of English, Volume 4: Africa, South and Southeast Asia* (pp. 610–623). Berlin/New York: Mouton de Gruyter.

Battarcharjee, D. S. (2009). *Lexical Innovation on the Internet-Neologisms in Blogs*. [Doctoral dissertation, University of Zurich].

Blommaert, J. (2010). *The Sociolinguistics of Globalization*. Cambridge: Cambridge University Press.

Bolton, K., Botha, W., & Kirkpatrick, A. (Eds.). (2020). *The Handbook of Asian Englishes*. USA: John Wiley & Sons Inc.

Brian, D. (2022, January 5). *How Many People Use Twitter in 2022? [New Twitter Stats]*. https://backlinko.com/twitter-users

Canagarajah, S. (2013). *Translingual Practice: Global Englishes and Cosmopolitan Relations.* Abingdon, UK: Routledge.

Chambers, J.K., & Schilling, N. (ed.). (2013.) *The Handbook of Language Variation and Change.* Oxford: Wiley-Blackwell.

Coats, S. (2016). Grammatical feature frequencies of English on Twitter in Finland. *English in Computer-mediated Communication: Variation, Representation, and Change,* 93:179-209.

Crothers, L. (2021). *Globalization and American Popular Culture*. Lanham: Rowman & Littlefield.

Dalzell, T. (2018). *The Routledge Dictionary of Modern American Slang and Unconventional English*. New York: Routledge.

Dovchin, S. (2018). *Language, Media and Globalization in the Periphery: The Linguascapes of Popular Music in Mongolia*. New York: Routledge.

Feltgen, Q., Fagard, B. & Nadal, J.P. (2017). Frequency patterns of semantic change: corpus-based evidence of a near-critical dynamics in language change. *Royal Society Open Science 4*(11): 170830.

Hamat, A., Embi, M. A., & Hassan, H. A. (2012). The use of social networking sites among Malaysian university students. *International Education Studies*, *5*(3), 56-66.

Ho A. I., Hashim, R. S., Jaludin, A., and Ismail, R. (2011). Enhancing opportunities for language use through web-based social networking. *International Conference on Social Science and Humanity 5*, 136-139.

Giddens, A. (1990). *The Consequences of Modernity*. Stanford: Stanford University Press.

Gilquin, G. (2018). American and/or British influence on L2 Englishes–Does context tip the scale (s). In S, C. Deshors (ed.), *Modeling World Englishes: Assessing the interplay of emancipation and globalization of ESL varieties* (pp. 187-216). Amsterdam: John Benjamins.

Gleason, B. (2018). Adolescents becoming feminist on Twitter: New literacies practices, commitments, and identity work. *Journal of Adolescent & Adult Literacy, 62*(3): 281-289.

Graddol, D. (2000). *The Future of English. A guide to forecasting the popularity of the English language in the 21st century*. (2nd ed.). London: The British Council.

Grieve, J., Nini, A. & Guo, D. (2017). Analyzing lexical emergence in Modern American English online. *English Language and Linguistics, 21*(1), 99–127.

Grieve, J., Montgomery, C., Nini, A., Murakami, A & Guo, D. (2019). Mapping lexical dialect variation in British English using Twitter. *Frontiers in Artificial Intelligence 2,* 1-18.

Grieve, J., Nini, A. & Guo, D. (2018). Mapping lexical innovation on American social media. *Journal of English Linguistics, 46*(4), 293–319.

Groff, C., Hollington, A., Hurst-Harosh, E., Nassenstein, N., Nortier, J., Pasch, H., & Yannuar, N. (Eds.). (2022). *Global perspectives on youth language practices*. Walter de Gruyter GmbH & Co KG.

Gulnazir, N., & Salehuddin, K. (2022). Investigating Lexical Variation and Change in Malaysian Twitter: A Conceptual Paper. *GEMA Online Journal of Language Studies*, *22*(4), 90-107. http://doi.org/10.17576/gema-2022-2204-06

Hashim, A. (2020). Malaysian English. In Bolton, K., Botha, W., & Kirkpatrick, A. (ed.). *The Handbook of Asian Englishes* (pp. 373-397). USA: John Wiley & Sons Inc.

Herring, Susan C., Androutsopoulos, J. (2015). Computer-mediated discourse 2.0. In Schiffrin, D., Tannen, D., & Hamilton, H. (ed). *The Handbook of Discourse Analysis (*pp. 127-151). Chichester, UK: John Wiley & Sons.

Huang, Y., Guo, D., Kasakoff, A., & Grieve, J. (2016). Understanding US regional linguistic variation with Twitter data analysis. *Computers, Environment and Urban Systems,* 59: 244-255

Izazi, Z. Z., & Tengku-Sepora, T. M. (2020). Slangs on social media: Variations among Malay language users on Twitter. *Pertanika Journal of Social Sciences & Humanities, 28*(1), 17-34.

Jensen, M. M. (2017). What can Google Trends data tell us about dialect labels: An exploratory study. *Globe: A Journal of Language, Culture and Communication, 5*, 48-76.

Kachru, B. (1985). Standards, codification and sociolinguistic realism: The English language in the Outer Circle. In Randolph Quirk & Henry G. Widdowson (eds.), *English in the world (*pp.11–30). Cambridge: Cambridge University Press.

Kashinathan, S., & Aziz, A. A. (2021). ESL learners' challenges in speaking English in Malaysian classroom. *International Journal of Academic Research in Progressive Educatioand Development, 10*(2), 983–991.

Khayrullina, R. & Fatkullina, F. (2021). The latest phraseology in internet discourse. In Arsenteva, E. (ed.), *The Discoursal Use of Phraseological Units* (pp. 264-283). Cambridge: Cambridge Scholars Publishing.

Kunalan, R., Mutty, B., & Francis, K. (2021). Lexical borrowing in Malaysian online newspaper editorials: A corpus-based approach. *International Journal of Education, Psychology and Counseling, 6*(40), 93-98.

Krazit, T. (2010, April 22). *The meaning of Google*. CNET. https://www.cnet.com/culture/the-meaning-of-google/

Lantz-Andersson, A. (2018). Language play in a second language: social media as contexts for emerging sociopragmatic competence. *Education and Information Technologies, 23*(2), 705-724.

Lee, J. S. (2016). "Everywhere you go, you see English!": Elderly women's perspectives on globalization and English. *Critical Inquiry in Language Studies, 13*(4), 319-350.

Mair, C. (2013). The World System of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide, 34*(3), 253–278.

Mahmud, F.N. & Salehuddin, K. (2023). How bilingual are Malaysian undergraduates? A snapshot of the different bilingual categories in Malaysia. *GEMA Online® Journal of Language Studies, 23*(2), 144-164. http://doi.org/10.17576/gema-2023-2302-08

Martin, E. (2020). World Englishes in the Media. In Cecil L. Nelson, Zoya G. Proshina, and Daniel R. Davis (ed.) *The Handbook of World Englishes (*pp. 597-615). John Wiley & Sons.

Mat Awal, N., Bakar, K. A., Abdul Rahman, A. N. C., & Abdullah, I. H. (2021). Representasi halal dan haram dalam wacana parlimen Malaysia. *GEMA Online Journal of Language Studies, 21*(4), 186-207. http://doi.org/10.17576/gema-2021-2104-10

Maybaum, R. (2013). Language change as a social process: Diffusion patterns of lexical innovations in Twitter. In *Annual Meeting of the Berkeley Linguistics Society, 39*(1): 152-166.

Meyer, Charles F. (2014). "Corpus-based and corpus-driven approaches to linguistic analysis: One and the same?" In Taavitsainen, Irma, Merja Kytö, Claudia Claridge, & Jeremy Smith (eds.), *Developments in English. Expanding Electronic Evidence.* Cambridge: Cambridge University Press, 14-28.

McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, Theory and Practice.* Cambridge: Cambridge University Press.

Moody A. (2020). The authority and authenticity of performative Englishes in popular culture. *World Englishes,* 1–21. https://doi.org/10.1111/weng.12486

Nevalainen T. (2015). Descriptive adequacy of the S-curve model in diachronic studies of language change. In *Can we predict linguistic change?* Helsinki, Finland: Varieng. http://www.helsinki.fi/ varieng/series/volumes/16/nevalainen/

Nor, Z. M., & Zamri, N. (2015). Lexical features of Malaysian English in a local English-language movie, Ah Lok Café. *Procedia-Social and Behavioral Sciences*, *208*, 282-299.

Nurmi, A. (1999). *A Social History of Periphrastic DO*. Helsinki: Société Néophilologique.

Pennycook, A. (2007). *Global Englishes and Transcultural Flows*. London, UK: Routledge.

Pillai, S., & Ong, L. T. (2018). English (es) in Malaysia. *Asian Englishes, 20*(2): 147-157.

Rosell-Aguilar, Fernando. (2018). Twitter as a Formal and Informal Language Learning Tool: From Potential to Evidence. In Rosell-Aguilar, F., Beaven, T. Gutiérrez, M.F. (ed.). *Innovative Language Teaching and Learning at University: Integrating Informal Learning into Formal Language Education* (pp. 99–106). https://research-publishing.net/manuscript?10.14705/rpnet. 2018.22.780.

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.

Salehuddin, K. (2022). Exploring three idealised cognitive models in *The Kampung Boy* and *Budak Kampung. 3L: The Southeast Asian Journal of English Language Studies, 28*(2), 140-159. http://doi.org/10.17576/3L-2022-2802-10

Salehuddin, K. & Winskel, H. (2016). Developmental milestone expectations, parenting styles, and self-construal of caregivers from Malay, Chinese and Indian backgrounds in Malaysia. Journal of Comparative Family Studies, XLVII (2 Spring Issue), 147-167.

Santoso, D. H., Aziz, J., Utari, P., & Kartono, D. T. (2020). Populism in new media: The online presidential campaign discourse in Indonesia. *Gema Online Journal of Language Studies, 20*(2), 115-133.

Sharma, B. K. (2012). Beyond social networking: Performing global Englishes in Facebook by college youth in Nepal. *Journal of Sociolinguistics, 16*(4), 483-509.

Simon, K. (2023, February 13). *Digital 2023: Malaysia.* https://datareportal.com/reports/digital-2023-malaysia

Sinclair, J. (2003). *Reading Concordances: An Introduction*. London: Longman Publishing Group.

Taboada, M. (2016). Sentiment analysis: an overview from linguistics. *Annual Review of Linguistics, 2*: 325–347.

Tagliamonte, S. A., & D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language*, *85*(1): 58-108.

Tankosić, A., & Dovchin, S. (2021). The impact of social media in the sociolinguistic practices of the peripheral post-socialist contexts. *International Journal of Multilingualism, 42*(6), 1-25.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.

Thirusanku, J., & Yunus, M. M. (2012). The many faces of Malaysian English. *International Scholarly Research Network* 1-14. https://doi.org/10.5402/2012/138928

Traugott, E. C. (2017). Semantic change. In *Oxford research encyclopedia of linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.323

Tseng, A., & Hinrichs, L. (2020). Mobility and the English Language. In Aarts, B., McMahon, A. & Hinrichs, L. (ed.) *The Handbook of English Linguistics* (pp. 637-652). UK: John Wiley & Sons Ltd.

Webb, T. (2018). Where the slang term 'lit' came from and how big brands have ruined it. https://www.abc.net.au/news/2018-01-17/where-the-slang-term-lit-came-from-and-how-big-brands-ruined-it/9335556 [30 November 2022].

Würschinger, Q. (2021). Social Networks of Lexical Innovation. Investigating the Social Dynamics of Diffusion of Neologisms on Twitter. *Frontiers in Artificial Intelligence.* https://doi.org/10.3389/frai.2021.648583

Yunus, M. M., Zakaria, S., & Suliman, A. (2019). The potential use of social media on Malaysian primary students to improve writing. *International Journal of Education and Practice, 7*(4), 450-458.

ABOUT THE AUTHORS

Noraishah Gulnazir is a lecturer at the Faculty of University Foundation Studies, HELP University. Her research interests include corpus linguistics, language variation and change, as well as globalisation of language.

Khazriyati Salehuddin (Ph.D) is an Associate Professor and a Psycholinguist at the Faculty of Social Sciences & Humanities, Universiti Kebangsaan Malaysia. She received her PhD in Psycholinguistics in 2010 from Western Sydney University. Khazriyati has led several research groups related to the area, including two national-level grants, namely Qur'anic Memorisation Techniques: A Psycholinguistic Module for Non-Arabic-Speaking Malay Speakers (FRGS) and Exploring the cognitive and perceptual processes in reading among Malaysian readers (ERGS). Khazriyati uses various methods in her research and has published several journal articles, books, and book chapters on the area. One of her works is published in South and Southeast Asian Psycholinguistics by Cambridge University Press and her book, Psikolinguistik: Penerokaan Minda Berlandaskan Bahasa, is published by UKM Press. She is currently the Head of Language & Cognition Research Cluster at her faculty. For more: goo.gl/rskBFV