Lexical Innovation Processes of Youth Netspeak on Malay Twitter Posts

Nur Nashatul Nasuha Nazman [a]
nashanazman94@gmail.com
Pahi Secondary School,
Kuala Krai, Kelantan, Malaysia

Su-Hie Ting [b]
shting@unimas.my
Faculty of Language and Communication,
Universiti Malaysia Sarawak, Malaysia

Kee-Man Chuah
kmchuah@unimas.my
Faculty of Language and Communication,
Universiti Malaysia Sarawak, Malaysia

ABSTRACT

The formation of neologisms in Malay has not been extensively studied and therefore has yet to be fully understood. The current study examined the lexical innovation processes in Twitter posts in Malay. A 25,679-word corpus was compiled from 2,000 tweets of 200 Malaysian Twitter users in their twenties. The textual analysis performed through AntConc software identified "*ni*" (this) and "*nk*" (want) as the top two neologisms, and these words constituted 45.8% of the corpus. Most of the neologisms are formed through clipping (65%) and pseudo-elliptical construction (26%), while blending (7%) and elongation (2%) are minor word formation processes. The clipping affects vowels (28%) more than consonants (9%), whether it is deletion, replacement, or addition. The pseudo-elliptical constructions involve letters of the alphabet and phonological adjustments to vowels and consonants. On vowel changes, the results show that there is a simplification of diphthongs to monophthongs, and a movement towards vowels in the middle position [e] or [o]. As for consonant changes, there is a shift in the place of articulation towards bilabial consonants [p], [m], and [w] and words containing [h] and [r] are likely to be dropped or replaced. The changes in the spelling are a result of changes in the spoken discourse. The findings suggest that exploring the phonological rules can explain pseudo-elliptical constructions, leading to a better framework on morphology and phonology to understand formation of neologisms in social media discourse.

Keywords: social media; Twitter; neologisms; non-standard words; Malay

INTRODUCTION

Social media discourse is a potent site to study and understand the evolution of language in the digital era. The instant messaging through digital media has brought about changes such as a more informal writing style and the use of emoticons and emojis for emotional expressiveness. For example, the acronym LOL (laughing out loud) appeared in internet slang and was first listed in The Oxford English Dictionary in March 2011 (Scholarly Community Encyclopedia, 2022). Since then, more words have been shortened in creative ways. One of the most used social network

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

eISSN: 2550-2131
ISSN: 1675-8021

microblogging services is Twitter. Twitter.com is largely text-based although attachments such as images and videos can be added (Gligorić, Anderson, & West, 2018; Jaidka, Zhou, & Lelkes, 2019). Twitter has popularised hashtags and newly coined words like "relfie" and "friendscrap" which refer to taking a selfie with a person one is close to and the action of adding all of one's friends as their own respectively (Yang, Ko, & Chung, 2019).

Numerous studies have been conducted on lexical innovations in social media discourse involving English words (Harared, 2018; Klymenko, 2019; Ljubotenska, 2018; Mutiah et al., 2021; Nevihayati, 2018; Suparta, Qomariana, & Rahayuni, 2016; Wu, Morstatter, & Liu, 2018). Lexical innovations in social media discourse include acronyms or abbreviations, phonetic replacements (e.g., "ur", Craig, 2003), onomatopoeic spelling (e.g., "hahaha", Barseghyan, 2013), clipping, and compounding (Barseghyan, 2013). In Malaysia, Gulnazir and Salehuddin's (2022) analysis of the use of lexical items "lit" (past tense of "light") and "on fleek" in tweets showed that the new meanings of "exciting" and "on point" respectively have spread from American English to Malaysian English.

Using Kiparsky's (1983) lexical theory of morphology and phonology, Nkhata and Jimaima (2020) found that Facebook and Whatsapp messages written by Zambian users showed non-standard English words formed mainly through acronymy/initialisation (40*%)* and clipping (30%). Our review of the literature showed that terms like abbreviation, acronym, and initialism are used interchangeably. Craig (2003) even had the inanities category for unclassifiable neologisms, indicating that word formation processes for netspeak have yet to be fully understood. Compared to English, less is known about social media discourse in Malay. However, research interest on it has grown in the last half decade (e.g., Mustafa et al., 2015; Ningsih, 2020; Paputungan, 2015; Sukma, 2014). Malay is the fourth most popular language used on Twitter (Hong, Convertino, & Chi, 2011). In Saloot, Idris, Aw, and Thorleuchter's (2014) Malay chat-style tweet corpus of 15 million words, 26 out of the top 40 most frequent terms are neologisms, which Saloot et al. (2014) referred to as out-of-vocabulary words, i.e., foreign words, abbreviations, or specific interjections (e.g., "*RT*", "*nak*" and "*ni*" to refer to "Retweet", "want", and "this" respectively).

Although neologisms have been extensively studied in English, there is a need to study neologisms in Malay because the spelling patterns differ in the two languages. Malay has the V, CV and CVC spelling patterns (Izazi & Tengku-Sepora, 2020). Here C refers to consonants and V refers to vowels. For example, "*bagaimana*" (how) shows the alternation of consonants and vowels. However, English words have a different spelling pattern. Analysis of seven-letter consonant patterns showed that CCVCCVC, CCVCVCV, CCVCCVV, CCVCVVC, and CVCCVCC as some of the common patterns (Thorpe, 2002). Furthermore, neologisms in English has creative abbreviation codes using letters of the alphabet and numbers for words such as "you" (U), "and" (N), "for" (4) and "to" (2). Combinations of letters of the alphabet and numbers like "b4" (before), "Y" (why) and "K" (okay) are used by Malaysian Twitter users (Chekima & Alfred, 2017; Omar et al., 2016).

In Malaysia, the number of Twitter users increased from three million in 2018 to 5.51 in 2021 (Statista, 2021). The volume of exchanges in social media messages written in Malay presents a conducive situation for the birth of new words and evolution of linguistic forms of the language.

Thus far, many of the studies on Malay social media discourse are descriptive, and provide lists of neologisms. Saloot et al. (2014) identified "RT" (a Twitter term), "*nak*" ("*hendak*" for want), "*ni*" (*"ini*" for this) and "*yg*" ("*yang*" for that is) as the top four words in a corpus of 321

Malaysian Twitter users who wrote in Malay. In Athirah et al.'s (2019) study on 25,000-word corpus comprising tweets and Facebook comments of Malaysians, "*mkn*" ("*makan*" for eat) topped the list. Athirah et al. (2019) found that "*nk*" (for "*hendak*" for want) ranked fourth while "*yg*" ("*yang*" for that) ranked second and "*je*" ("*sahaja*" for only) ranked third.

Some researchers have gone further to deduce word formation rules for neologisms; however, there is a problem of unclear and overlapping categories. For example, some researchers did not operationally define terms such as conversion (Paputungan, 2015) and derivation (Nisa, 2016), and this leaves the terms open to the interpretation of readers. Readers have to use examples to deduce that the term "conversion" in Suparta et al. (2016) refers to noun/verb word class change because operational definitions are not provided. Athirah et al.'s (2019) list of 16 word formation rules for neologisms in the Malay social media discourse is by far the most detailed; yet, they cannot explain all the newly-formed words in their database because of the unclear boundaries. For example, for a word ending with "a" (e.g., "*apa*" for what), the rule is to replace it with "e" ("*ape*"). While this rule can explain why "*siapa*" (who) is written as "*siape*", it cannot explain why "*juga*" (also) is written as "*gak*" and not as "*juge*". Part of the problem is that the derivation of word formation rules is not underpinned by a clear theoretical framework, which makes it difficult to compare findings across studies and settings. For the field to advance especially in assisting the process of formulating algorithms to normalise various newly coined words, it is important to work descriptively on a clearly limited set of processes at work in the language.

The present study examined the lexical innovation processes in Twitter posts written in Malay. The specific objectives were to: (1) identify the top 100 neologisms in Malay tweets; (2) determine the percentage of neologisms in Malay tweets which carry new meanings; (3) describe morphological changes for the form of the top 100 neologisms in Malay tweets; and (4) deduce the lexical innovation processes of the top 100 neologisms in Malay tweets based on a theory of lexical morphology and phonology.

THEORETICAL FRAMEWORK OF STUDY

This study adapted Kiparsky's (1983) theory of lexical morphology and phonology as the theoretical framework used for analysing word formation processes for words with non-standard spelling. Lexical morphology is a branch of morphology which investigates the processes by which words are formed (Nkhata & Jimaima, 2020). Kiparsky's (1983) theory of lexical morphology and phonology was based on Pesetsky's (1979) proposed grammatical model where the word structure is mapped onto the logical form (as cited in Pesetsky, 1985).

Kiparsky (1983) applied a level-ordered morphology for underived lexical items in the English lexicon, whereby the phonological rules apply within the lexicon after each step in the morphological derivation of a word (Figure 1). The left side of the figure shows the word formation processes whereas the right side of the figure shows the rules of lexical phonology. At level 1, the primary morphology involves adding a suffix (i) -an to a stem (e.g., from Mendel to Mendelian). At level 2, the secondary morphology involves adding -ism (e.g., Mendelianism). Level 2 includes compounding. This level-ordering predicts that there are no words at all in *-ismian (e.g. *Mendelismian). Kiparsky (1983, p. 5) stated that "-ism does not participate in the assignment of word stress and is not followed by primary suffixes because it is added at level 2, where word stress does not apply and primary suffixes are not available". The rules of lexical phonology which apply within the lexicon are word stress and trisyllabic shortening at level 1 and compound stress

at level 2. After words are put together by the syntax, the rules of postlexical phonology applies (e.g., sentence stress and aspiration).

Using Kiparsky's (1983) lexical approach to morphology and empirical studies (Mworia, 2015; Onyedium, 2012), Nkhata and Jimaima (2020) identified the following lexical innovations (which were referred to as morphological processes in their publication) in Facebook discourse: abbreviations, acronyms, affixation (derivation and inflection), blending, borrowing, clipping, compounding, conversion, initialism, semantic extension, and coinage. They introduced a category named "pseudo-elliptical construction" for new words with unclear morphological processes. However, Nkhata and Jimaima (2020) did not classify the processes according to levels as in Kiparsky's (1983) theory of lexical morphology and phonology; they generally did not deal with postlexical phonology. In this paper, we show that the application of the phonological aspect of Kiparsky's (1983) theory of lexical morphology and phonology can explain productive word formation rules in Malay social media discourse.

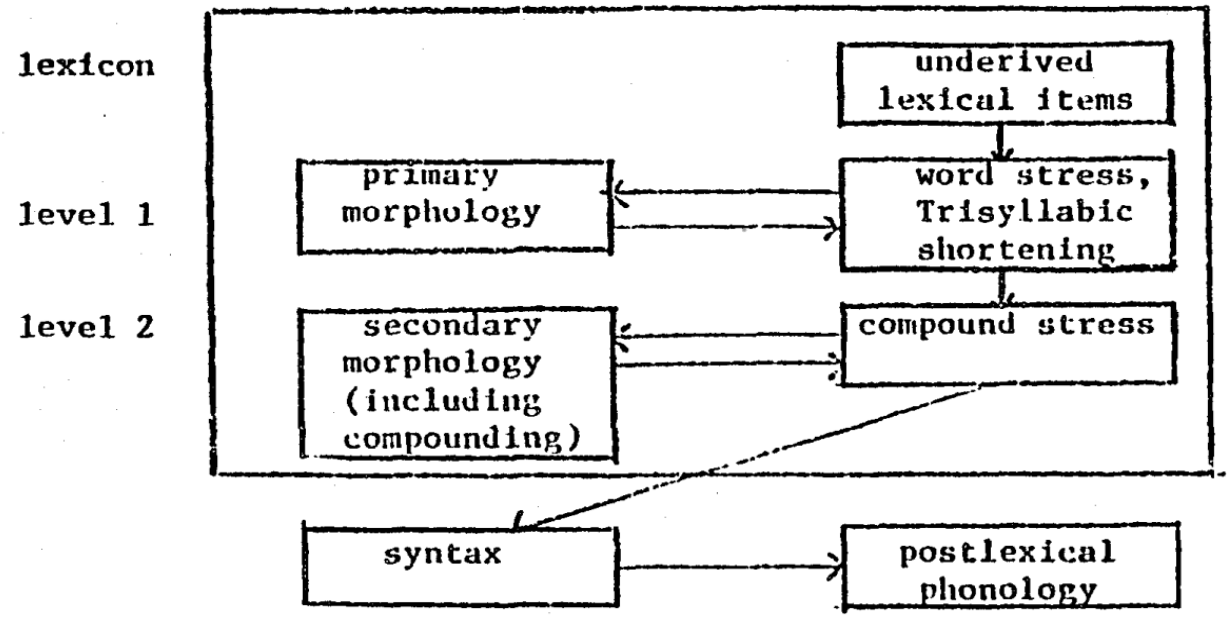


FIGURE 1. Word-formation processes in Kiparsky's (1983, p.4) theory of lexical morphology and phonology

METHOD

A descriptive research design was chosen for this study on neologisms in Malay tweets to identify the characteristics, frequencies, and categories of lexical innovations of social media language that have not been extensively investigated.

PARTICIPANTS

The data were tweets written in Malay by Malaysian Twitter users in their twenties. The Twitter users were in the social circle of the first author, which is why their age is known. The participants were 200 active Twitter users with a public profile. They fulfilled the definition of active Twitter users, which is, they follow at least 30 accounts and are followed in return by at least one-third of

eISSN: 2550-2131
ISSN: 1675-8021

those accounts (Twitter, 2020). The participants selected were those who had posted more than 30 Malay tweets to ensure there would be enough content in the target language for analysis. A total of 2,000 tweets were collected from 200 Twitter users (10 tweet episode per user) which results in a 25,679-word corpus.

INSTRUMENT

For the word formation processes, Nkhata and Jimaima's (2020) adaptation of Kiparsky's (1983) theory of lexical morphology and phonology was employed (Table 1). For the analysis of meaning change, the first author referred to an established Malay dictionary, Kamus Dewan (2007). For analysis of the form of the neologisms, the framework can be seen in Table 2.

TABLE 1. Nkhata and Jimaima's (2020, pp. 78-87) identification of word formation rules that underlie morphological processes

Morphological processes	**Definition**	**Word Formation Rules**	**Examples**
Acronym	Acronym involves inclusion of all word initials regardless of whether a given word is grammatical (e.g., articles, prepositions or affixes) or lexical (e.g., nouns, verbs).	[+ initials]	BTW for By The Way
Affixation	Affixation involves adding affixes to already existing words and includes prefixes, infixes and suffixes added to the beginning, middle and end of words respectively.	Prefix [un-] + [+Acronym]	unUN for un United Nations
		[Self] Noun + Inflection [-ie]	Selfie for Self + ie
Clipping	Clipping involves cutting off part of a word and still being a member of the same for class and maintaining the meaning of the whole word.	[+ truncation]	Hun for Honey
Compounding*	Compounding involves the combination of two distinct words into one word, and the spelling of the original words are still retained.	[+ comp]	Beetweet for Bee + tweet**
Blending*	Blending involves the merging of parts of two or more words into one word and the original words cannot be easily separated from the new word.	[+ blend] [+ elliptical]	Twitosphere** for Twitter + Atmosphere
Reduplication	Reduplication involves repetition of phonemes.	[+ redup]	Hehehehehe for laughter
Pseudo-elliptical construction	Pseudo-elliptical construction involves word omissions to form phrases that do not show any structural features or constituents for phrases. Includes coinage of new words that flout established word formation processes.	[+ elliptical]	Gonna for going to

Notes:
**With reference to Renner (2019), we have redefined compounding (e.g., appletree) and blending (e.g., smoke+fog=smog) for clarity.*
***Examples from Suparta et al. (2016).*

eISSN: 2550-2131
ISSN: 1675-8021

DATA COLLECTION AND ANALYSIS PROCEDURES

The first author contacted her Twitter contacts, informed them of the study and requested voluntary participation. The participants were told that those who qualify are those who had at least one episode of tweeting in a month, with each episode comprising about four to five tweets. The participants were asked to submit 10 tweets each, and altogether the 200 Twitter users submitted a total of 2,000 tweets. The corpus size was 25,679 words.

The data were filtered by removing extra characters such as hashtags or symbols before inputting into the AntConc textual analysis software. AntConc is a multi-platform, multi-purpose corpus analysis toolkit, designed specifically for use in the classroom. It hosts a comprehensive set of tools including a powerful concordance, word and keyword frequency generators, tools for cluster and lexical bundle analysis, and a word distribution plot (Anthony, 2005). The AntConc software ran through the compiled texts and generated a list of words and the frequency counts based on their stem forms.

Based on the listing of words generated by the AntConc software based on frequency, the top 100 words with spelling that differ from what is in Kamus Dewan (2007) were identified for further textual analysis. The choice of the top 100 words is more than Saloot et al. (2014), who used the top 40 most frequently used Malay words in their 15-million-word corpus of Twitter messages. However, Athirah et al. (2019) also used the top 100 words from their corpus of 25,000 words for analysis.

RESULTS

FREQUENTLY USED NEOLOGISMS IN TWEETS BY TWITTER USERS

The top two neologisms in the 2,000 Malay tweets were "*ni*" (this) and "*nk*" (want), shortened from "*ini*" and "*nak*" respectively. These two words account for 30.63% and 3.45% of the cumulative frequency of the top 100 neologisms in Malay tweets (n=2,955).

The analysis of the top 100 neologisms in Malay tweets showed that not all new words carry new meanings because many of them were only shortened versions of words in Kamus Dewan (2007). Out of the 100 neologisms, 76 words retained their original meaning, 18 carried new meanings whereas six had either the same or different meanings depending on context of use. Table 2 shows the top 10 words out of the top 100 neologisms in Malay tweets.

TABLE 2. Examples of neologisms with changed or retained meanings

	Neologism	Standard form	Word class	Meaning in the Dictionary	Retention of dictionary meaning
1.	*Ni*	*Ini*	Determiner	*Utk menunjuk waktu, tempat atau benda yg dekat* (to show time, destination and nearby object)	Yes
2.	*Nk*	*Hendak*	Verb	*Ingin mempunyai, mendapat atau memperoleh sesuatu* (to have, to get)	Yes
3.	*Nie*	*Ini*	Determiner	*Utk menunjuk waktu, tempat atau benda yg dekat* (to show time, destination and nearby object)	Yes

4.	*Ak*	*Aku*	Pronoun	*kata ganti nama diri pertama* (first person pronoun)	Yes
5.	*Ja*	*Sahaja*	Adverb/adjective	*Cuma, hanya* (only)	Yes
6.	*Da*	*Sudah*	Adverb/adjective	*Telah siap sedia; selesai; habis; tamat* (ready, done, the end)	Yes
7.	*X*	Tidak	Determiner	tiada; bukan (no)	No. X in standard use is just a letter of an alphabet but in the tweets, X is used with the meaning of negation at the phrase level.
8.	*Dok*	*Duduk*	Verb	*Meletakkan badan bertumpu pada punggung, tinggal, tempat* (sit down, stay)	Yes. "Dok" in its standard form means "dockyard" but when it is used as a shorter version of "duduk", there is no change in meaning.
9.	*Duk*	*Duduk*	Verb	*Meletakkan badan bertumpu pada punggung, tinggal, tempat* (sit down, stay)	Yes when it is used as a shorter version of "duduk". No when it is used to ask a question. For example, "*duk buat apa?*" is formed from three words: "*duk buat apa*" which means "what are you doing?" Admittedly, "duk" here could be an accidental homonym.
10.	*Lak*	*Pula*	Adverb	*Berulanglagi; sekalilagi*(again/well/neither)	Yes when it is a shorter version of "pula". No when it is used as a question tag. For example, "*mana lak*?" (where is it?) Admittedly, "lak" here could be an accidental homonym.

Note: The word classes are based on the Malay dictionary, Kamus Dewan (2007).

The two top ranked words, "*ini*" and "*nak*", retained the dictionary meanings. "*Ni*" is also spelt as "*nie*" but this spelling is not as popular because it ranks fourth. Other neologisms that retain the meaning of original standard words are "*ak*" for "*aku*" (I), "*ja*" for "*sahaja*" (only) and "*da*" for "*sudah*" (already).

Table 2 shows an example of words that are used with different meanings in Malay tweets. X is the 24th letter of an alphabet but in the tweets, X is used with the meaning of negation. "*Dok*" in its standard form means a shipyard but Twitter users use "*dok*" to mean "*duduk*" (sit down), and there is no change in meaning. The use of "p" with different meaning is shown in Table 2. In standard usage, "p" is the 16th letter of the alphabet but in Malay tweets, it carries the meaning of "go"in the Penang Malay dialect ("pi" is a shortened version of "pergi"). Hence, it is not a surprise that "p" is used for convenience to speed up writing of tweets. As for "*lau*", it retains the meaning

if it is used to replace "*kalau*" but it is also a Malay word for the Moth orchid species (*Phalaenopsis amabilis*) and a Chinese surname.

Six of the top 100 neologisms may be used with the same or different meaning, depending on their contexts. For instance, "*lak*", when used as a shortened version of "*pula*" (repetition), retains the dictionary meaning but when it is used for emphasis in a question, its meaning depends on the context. This is based on general observations of the authors who are Malaysian speakers of Malay. For example, "*mana lak semua orang*?" (where is everyone?). Admittedly, "lak" here could be an accidental homonym. Another neologism that takes on a different meaning is "*duk*", an abbreviation of "duduk" (sit down/stay). When it is used to ask a question, it does not mean sit. For example, "*duk buat apa?*" is formed from three words, which means "what are you doing?" Admittedly, "*duk*" here could be an accidental homonym.

The analysis shows that a small proportion of neologisms in Malay tweets have different meanings from standard words, but Twitter users with a shared sociocultural background can figure out the meaning from the context.

LEXICAL INNOVATIONS IN THE FORM OF NEOLOGISMS IN MALAY TWEETS

Table 3 shows that there are more lexical innovations involving vowels (28%) than consonants (9%), whether it is deletion, replacement, or addition. There is a greater need to retain consonants because they form the shape of the word and their presence makes the words unrecognisable. Syllable deletion with or without additional changes accounts for 19% of the changes while multiple deletion leaving consonants that accounts for 18% of the top 100 neologisms used by Malaysian Twitter users. One quarter of words (i.e., 25 words) cannot be classified based on the form.

Table 3 shows that 53% of the top 100 neologisms are formed through four types of deletion. Vowel deletion usually results in words that cannot be pronounced (e.g., "*smpai*" from "*sampai*" meaning "arrive"). The new spelling has no autonomous pronunciation. Hence, no new word is formed (as in the case of true clipping like "sudah" (already) which is clipped to "dah"). Vowel deletion produces a shortened graphic variant, an abbreviation to speed up typing of tweets. However, when consonants are deleted, the words can still be pronounced (e.g., "*tau"* from "tahu" meaning "know"). Syllable deletion usually results in words that can be pronounced, for example, "*fon"* from "*telefon*" (telephone).

When there is a vowel replacement, the new word sounds similar, particularly when the neutral schwa is used (e.g., "*keje*" from "*kerja*" for work or "*tido"* for "*tidur"* for sleep). As for consonant replacement, the pronunciation of the original and new words may differ (e.g., "*aq"* for "*aku"*, meaning I). Interestingly, consonants are sometimes added but not vowels. The letter "k" is added at the end of words such as "*pulak*" and "*jugak*". However, there are only some instances like this because shorter words can save time in typing messages.

We have attempted to formulate mutually exclusive categories in our typology based on the form of neologisms in Malay tweets and ensured that the categories are parallel, and not a mixture of descriptive categories and processes.

eISSN: 2550-2131
ISSN: 1675-8021

TABLE 3. *Frequencies of different categories of lexical innovations in the form of neologisms in Malay tweets*

Category	Definition	Example	Frequency
1. Vowel deletion	A vowel in the word is deleted to make it shorter.	"*smpai"* from "*sampai"*	20
2. Consonant deletion	A consonant in the middle or end of words is deleted to make it shorter and the new words can still be pronounced.	"*tau"* from "*tahu"*	6
3. Syllable deletion	The syllable deletion may be accompanied by a vowel change, a consonant change, a vowel addition, or consonant addition.	"*fon"* from "*telefon"* Syllable deletion + vowel change: "*de"*from "*ada"* Syllable deletion + consonant change: "*aq"*from "*aku"* Syllable deletion + consonant addition: "*gak"*from "*juga"*	9 6 1 3
4. Vowel change	An original vowel is replaced by another vowel.	"*ade"*from "*ada"*	8
5. Consonant change	A consonant is replaced by another consonant. Sometimes the new word cannot be pronounced.	"*pasal"*from "*fasal"*	1
6. Consonant addition	A consonant is added, usually at the end of words.	"*pulak*" from "*pula*"	2
7. Consonant frame	One or more letters of a word is deleted leaving the consonants which give shape to the word	"*tgk*" from "*tengok*"	18
8. Unclassifiable	Unclear patterns of word formation	"*keje*" from "*kerja*"	26
Total			100

WORD FORMATION PROCESSES UNDERLYING LEXICAL INNOVATIONS

Table 4 shows that most of the neologisms were formed through clipping (65%) and pseudo-elliptical construction (26%), with blending and elongation accounting for only 9% of the words.

TABLE 4. Frequency of morphological processes for top 100 neologisms in Malay *tweets analysed in this study*

Morphological process	Word formation rules	Frequency
Clipping	[+truncation] [truncation], [+ elongation]*	65
Blending	[+truncation], [+ blend]	7
Elongation	[+ consonant]	2
Pseudo-elliptical construction	[+truncation], [+ elliptical]	26
Total		100

*Note: *Subsequent step which occurs for some truncated words*

Clipping involves removal of parts of a word, and the applicable word formation rule is [+ truncation]. Most of the clipping involved the front syllable of a word and the resulting non-standard word can still be pronounced. For example, the first syllable of the adverb "*sudah*" (already) is clipped leaving "*dah*" or "*da*". The conjunction "*tetapi*" (but) is truncated to "*p*" which is similar in pronunciation to "*pi*". The clipping also frequently involves several vowels and consonants, leaving a consonant frame which cannot be pronounced. For example, "*tengok"* (see) is truncated to "*tgk"* or "*hendak*" (want) to "*nk*". "*Hendak*" is not truncated to "*hdk*" because "*nk*" is closer to "*nak*", an expression that is frequently used in daily conversations. Our analysis indicated that the truncation usually leaves the initial, middle, and last consonant to frame the word. Occasionally, the second vowel may be retained such as "*bila*" (when) which is clipped to "*ble*", "*kalau*" (if) which is clipped to "*klau*" or "*klu*", and "*sampai*" (arrive) which is clipped to "*smpai*". It is probable that a consonant frame without any vowels is too elliptical to be understood: "bl", "kl" and "smp" respectively. In the Malaysian context, "kl" is ambiguous because it is an acronym for the capital, Kuala Lumpur.

The analysis revealed the frequent deletion of "h" when it is the final consonant and the pronunciation of the word is not affected much. For example, "h" is clipped from "*boleh*" (can) to form "*bole*" and to "*tahu*" (know) to form "*tau*".

There were nine instances of clipping accompanied by the addition of a consonant (n=6) or a vowel (n=3) at the end. An example of a truncated word with an added consonant is "*gak*". The adverb "*juga*" is truncated to "*ga*" and the letter k is added. Another example is "*tuh*" formed from "*itu*". An example of a clipped word with an added vowel is "*ini*", which is truncated to "*ni*" and the letter "e" is added. "*Nie*" is not idiosyncratic use because it ranks fourth among the top 100 neologisms with a frequency of 73.

Next, the formation of seven neologisms through blending is described. Blending involves the merging of parts of two or more words into one word and the original words cannot be easily separated from the new word. The word formation rules that are applicable are [truncation] + [blend]. First, the phrase "*selepasitu*" ("*selepas itu*" meaning after that) is combined into "*pastu*" ("*pas tu*"). The last syllable of the first word ("*pas*") is combined with the last syllable of the second word ("*tu*") to form a new word that can be pronounced. All other syllables are clipped, that is, the initial syllables of the first word "*sele*" in "*selepas*" and "i" in the second word "*itu*". The meanings of the root words are retained. Blending also explains the formation of "*tujela*" from "*itusahajalah*" (that's all) and "*pehal*" from "*apahal*" (what's up?). The parts that are combined are underlined. These are creative shortening of words, like "*arini*" (today) from "*hariini*", where *ari+ni=arini*. Similarly, "*tak yah*" (not necessary) is formed from "*tidakpayah*", where *tak+yah=takyah*. "*Takyah*" is not strictly a blend because the word "*tak*" is a shortened version of "*tidak*" but there is still a merging of "*tak*" and "*yah*". "*Tak*" is a shortened version of "*tidak*" commonly used in speech but not in tweets because Twitter users use X for negation at phrase and sentence level. Similarly, "*cenni*" (like this) is formed through a similar blending process from "*macamini*" and undergoes a spelling change: *cam+ni=camni* →*cenni*. (The spelling change will be addressed in the results on pseudo-elliptical construction later). There is a consistent pattern in that the syllables retained are usually the last syllables of the words. The exception is when the blended word is an existing word with a different meaning. For example, "*apasal*" (what's the matter?) is formed from "*apafasal*", and the contracting of words is slightly different: *apa+sal=apasal*. If only the last syllables of the words are retained, then the resulting word is "*pasal*". "*Pasal*" is another Malay word that carries the meaning of issue. What can be concluded

eISSN: 2550-2131
ISSN: 1675-8021

is that blending in Malay social media discourse usually involves merging the last syllables of the original words to form new words.

This study found two out of the eight formal operations involved in word-formation processes of English identified by Renner (2020), which are clipping and blending. The six other morphological operations were not found in the neologisms used by Malay Twitter users, namely, prefixation, suffixation, compounding, morphostasis, stress shift, and desuffixation. One of the less common word-formation processes identified in this study is word elongation, which is the process of adding one or more letters to a word in order to create a longer version of that word. Elongating words is rare and can be considered as a minor word-formation type (Bauer, 2006) particularly in the context of social media texts where words are often shortened to increase efficiency. Nevertheless, word elongation is present and can be used to add a sense of informality to a message that can help to convey an emotional tone or add emphasis on a word (Baldwin & Li, 2015; Tchokni, Seaghdha, & Quercia, 2014).

In the corpus of Malay tweets analysed in this study, out of 100 neologisms, there are two instances of elongation; the other 98% of the words involve shortening. The word "*pulak*" is formed from "*pula*" (the adverb well) through a process of adding the consonant k [+ consonant]. "*Pulak*" ranks 74 and was used 11 times by the Twitter users but the shortened version "*lak*" is far more popular as it is the top 10 neologisms. The process is similar for the formation of "*nasik*" (rice) from "*nasi*". The addition of the letter k when the word "*nasi*" is typed in tweets marks the informal context of communication as it is reflective of how the words are spoken in informal interactions. As found by Rusli et al. (2022), Malay language learners have the habit of adding the consonant k to certain Malay words even in formal writing as those words are spelt in such a way in daily communication. Another example of consonant addition is the use of "r" in "*sumer*" for "*sama*" (meaning the same). Word elongation is not suffixation like how truncated words (e.g., "*mne*") are formed from clipping of vowels in "*mana*" to form "*mn*", followed by addition of a neutral vowel. This is because suffixation means that the final consonant that is added must carry an independent and stable meaning. Elongated words are regarded as an anomaly considering the quick pace of sending social media messages. In social media discourse, words are conventionally shortened but these two instances show elongation of words to mark informality and emotional expressiveness (Baldwin & Li, 2015; Parkins, 2012). Izazi and Tengku-Sepora (2020) found word elongation in Malay tweets for emphasis but such instances were not found in the present study (e.g., "*apa*" elongated to "*apaaaa*" (what); "*tapi*" elongated to "*tapiiii*" (but); "*wah*" elongated to "*waaa*").

Finally, our analysis of the phonological aspect of the formation of the 26 neologisms put in the pseudo-elliptical construction category shows that the loose [+ elliptical] word formation rule can, in fact, be explained. According to Kiparsky's (1983) theory of lexical morphology and phonology, at each level the output of the word formation process is subjected to a set of phonological rules. Our analysis uncovered pseudo-elliptical construction dealing with negation, letters of the alphabet, vowel change, and consonant change.

The use of X (in both upper- and lowercase) for negation is not new. In Malay tweets, "x" can replace "*tidak*", "*tak*", "*bukan*" and "*tiada*". The applicable word transformation rule is [+ elliptical] because "*x*" is not taken from the original words. Some Twitter users use "*xde*", in which case the applicable word transformation rule is [+ elliptical]. "*Xde*" is from "*tidak ada*" or "*tak ada*" and "*tak*" is written as "x". The vowel in the second syllable of "*ada*" is changed to the neutral vowel but spelt with an "e". The Twitter users are not linguistically trained and therefore

do not know how to use schwa and the keyword also does not permit use of the upside-down e, i.e., "ə".

The use of letters of the alphabet to form new words is not new either. In the corpus of Malay tweets, both "p" and "g" are used to represent "*pergi*" (go). "p" for "*pergi*" is ranked at number 53 while "g" for "*pergi*" is ranked at 57. In the top 100 neologisms in Malay tweets, "pg" for "*pergi*" was absent. In English, there are even more shortenings of words to letters of an alphabet such as "u", "i", "n", and "y" for "you", "I", "and" and "why respectively.

The more interesting pseudo-elliptical construction involving vowel changes is linked to phonological adjustments. Figure 2 shows the three vowel changes, namely, (1) changing [u] to [o], (2) changing [a]to the neutral vowel [ə], and (3) changing [i] to [e]. The fourth pseudo-vowel change is a matter of spelling: writing [ia] as y. The changes were used to deduce rules that are applied to explain the changes in a systematic way.

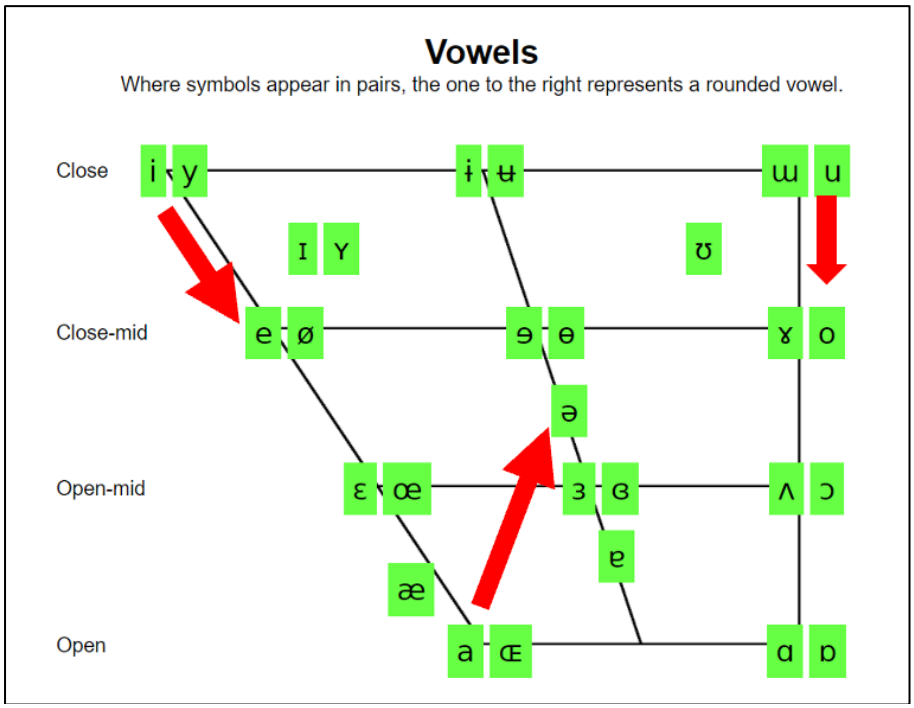


FIGURE 2. Vowel changes in neologisms formed through pseudo-elliptical construction
(Source: Vowel chart, https://www.ipachart.com/)

Firstly, the back vowel [u] is replaced by [o], another back vowel. The change is from a close to a close mid position. The [u] to [o] rule explains the change from "*pun*" (even) to "*pon*" or "*pom*". The [u] to [o] change reflects how the word is actually uttered as in reality, "*pun*", for example, is hardly manifested as [pun] in the spoken discourse, but as [pon]. For "*pom*", the consonant change will be explained in a later part of the results. This rule, in fact, also explains the shortening of words with the diphthong [au] because the [u] is changed to [o] in "*engkau*" (you) which is shortened to "*ko*". Here, the diphthong is simplified to a monophthong. The [u] to [o] rule also explains the change from "*tidur*" (sleep) to "*tido*". The same goes with other words; for example, "*ada*" is normally pronounced as [adə] than [ada]. The vowel shifts in the "tweets" most likely reflect the spoken discourse.

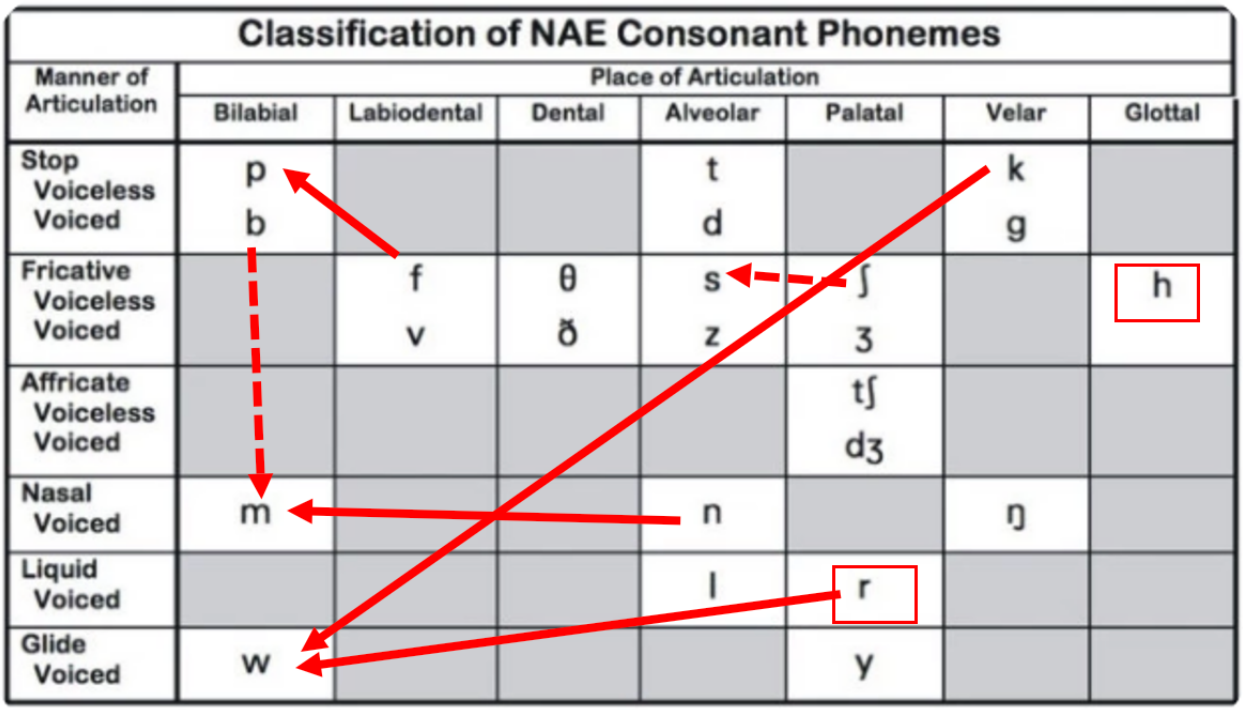
eISSN: 2550-2131
ISSN: 1675-8021

Secondly, the open front vowel [a] is replaced by the neutral vowel, spelt as an "e" by the Twitter users. This [a] to [ə] rule explains the vowel change in words with the [a] vowel such as "*ade*" which is formed from "*ada*" (there is), "*Xde*" from "*tiada*" (none), "*spe*" from "*siapa*" (who), "*pe*" from "*apa*" (what), "*ble*" from "*bila*" (when) and "*kite*" from "*kita*". The [a] to [e] rule also explains why "*semua*" (all) is written as "*sumer*" because the diphthong [ua] is simplified to [ə].

Thirdly, the front close vowel [i] is replaced by another front vowel but in mid-position [e]. This [i] to [e] rule explains why "*kahwin*" is written as "*kawen*" (marry) and why "*habis*" is changed to "*abes*" (finished).

Lastly, the fourth pseudo-vowel change affects the pronoun "*dia*" which is variously written as "*die*", "dye" and "dy". The formation of "die" can be explained by the [a] to [ə] rule, and this spelling is not popular; it ranks at 82. Writing [i] as "*y*" is a matter of spelling but, interestingly, "dye" and "dy" have gained traction among the Twitter users because they rank 67 and 68 respectively among the neologisms in Malay tweets.

The analysis of neologisms with vowel changes formed through pseudo-elliptical construction shows two phonological adjustments. There is a simplification of diphthongs to monophthongs, or monophthongisation. There is also a movement of vowels towards the middle position, whether it is from a close front [i] or back [u] position or from an open position [a]. This explains why many shortened words have the vowels [e] or [o]. Note, however, that both [e] and [ə] are spelt with an "e" by Malaysian Twitter users in their twenties. The phonological adjustments that apply to the truncated words is an interesting finding because tweets are meant to be read but the phonological adjustments to vowels suggest that the users actually imagine how they would say the words when they compose or read their tweets.



Note: Dotted arrows show consonant deletion, solid arrows show replacements, and squares show consonants usually dropped

FIGURE 3. Consonant changes in neologisms formed through pseudo-elliptical construction
(Source: Consonant sounds, https://www.pinterest.com/pin/319966748531016245/)

eISSN: 2550-2131
ISSN: 1675-8021

Next, the pseudo-elliptical constructions involving consonant changes linked to phonological adjustments are explained. Figure 3 shows the five consonant changes, namely, (1) consonant deletion involving [b] and [ʃ], (2) dropping [r], (3) changing [f] to [p], (4) changing [n] to [m], and (5) changing consonants to [w]. Firstly, consonant deletion involving [b] and [ʃ] makes the new words easier to pronounce. For example, "*ambil*" (take) is truncated to "*amik*". Both are bilabials but the voiced nasal [m]. The alveolar liquid [l] is also changed to the velar stop [k]. Using "k" in the end position in words such as "*nasik*" and "*pulak*" is consistent with the pattern of words formed through suffixation (with or without clipping) where a "k" is added (described earlier). The word "*asyik*" (pronounced as [aʃik]) is truncated to "*asik*" because the voiceless palatal fricative [ʃ] is more difficult to pronounce than the voiceless alveolar fricative [s]. Lutfiyana's (2020) study on Indonesian students in pronouncing English sounds showed that some students had difficulties with these sounds ([ʃ], [ʒ], [ʧ], [ʤ]) and tried to substitute them with the existing sounds in their native language.

Secondly, the dropping of [r] to form neologisms happens to two words with [r] in the middle syllable. "*Kerja*" becomes "*keje*" through a process of dropping [r] and changing the vowel [a] to the neutral vowel. The dropping of [r] is also seen in "*cerita*" (story) which is written as "*cter*". The middle syllable with [r] is dropped because the syllable is unstressed. The vowel [a] is changed to the neutral vowel and written as "*ter*" but can also be written as "*te*". In another study, Chekima and Alfred (2017) found that "*cerita*" is shortened to "*cite*". Izazi and Tengku-Sepora (2020) has "*cterrr*". Although the spelling of the truncated words is slightly different, the dropping of the middle syllable with [r] is also evident in these two studies. The dropping of [r] is not idiosyncratic use because "*keje*" ranks 55 in frequency among the neologisms with 14 instances. The frequency of "*cter*" is lower at the rank of 100 with nine instances in the dataset. However, when [r] is in the last syllable and crucial to show the shape the shortened words, they are retained like in "*bru*" (for "*baharu*" meaning new), "*sorg*" (for "*seorang*" meaning a person) and "*skrg*" (for "*sekarang*" meaning now). Chekima and Alfred's (2017) dataset also has "*sorg*" for "*seorang*".

Thirdly, the voiceless labiodental fricative [f] is replaced by the voiceless bilabial stop [p] in the word. An interesting example is the word "*fasal*" (issue) which is clipped to produce "*pasal*", and ranks 46 in the list, indicating a rather frequent usage by Twitter users. The consonant change is reflective of the interchangeable use of the words in daily conversations. The pseudo-elliptical construction involving the [f] to [p] change shows a change in the place of articulation from the labio-dental to the bilabial position. Phonologically, another change is the place of articulation. The first is a fricative and the second is a plosive.

Fourthly, the [n] to [m] change is described. In the corpus, "*pun*" (even) written as "*pom*" ranks 92 with 10 instances. Both are voiced nasals but the place of articulation changes from the alveolar for [n] to the bilabial position for [m].

Lastly, the analysis revealed a gravitation towards [w] in the formation of neologisms, and this explains the pseudo-elliptical construction of three words. The noun "*orang*" (people) which is written as "*owg*" and this word ranks 89 in frequency. The shortened form cannot be pronounced, and here the <w> in "owg" refers to the written form. Here, the palatal liquid [r] is replaced by the bilabial glide [w]. The verb "*tahu*" is spelt as "*taw"*, after the consonant as "h*"* is clipped. The pronoun "*aku*" (I) is spelt as "*aq*" and [q] is pronounced as [kw]. There seems to be a slurring in pronouncing "*orang*" and "*tahu*" as "*owang*" and "*taw*" respectively, and the use of [w] makes the pronunciation of the two syllables less distinctive, like how the words would be said in rapid speech.

eISSN: 2550-2131
ISSN: 1675-8021

The five categories of pseudo-elliptical constructions involving consonant change show two patterns in phonological adjustments. There is a change in the place of articulation but the voicing quality of the consonants is retained in [ʃ] to [s], [f] to [p], [b] to [m] and [n] to [m]. The arrows in Figure 3 clearly show that there is a shift in place of articulation towards the bilabial consonants. This means that neologisms are more likely to have bilabial consonants ([p], [m], and [w]) than those articulated further back in the mouth. In other words, if the original words contain [h] and [r] in the spelling, they are likely to be dropped or replaced. If the original words contain [k], [ʃ], and [n], they are likely to be replaced.

The results show that by delving into phonological aspect of Kiparsky's (1983) theory of lexical morphology and phonology, we have successfully formulated some word formation rules that can account for the 26 neologisms with pseudo-elliptical constructions, and in general for the dataset of Malay tweets analysed.

DISCUSSION

The study on youth netspeak in Twitter posts written in Malay shows that lexical innovations underlying neologisms are largely rule-governed. The three key findings of this study are discussed. Firstly, the five top ranked words comprising determiners and general verbs are also found to be high in frequency in other studies on Malay social media discourse. However, there seems to be an indication of change over time (about eight years) based on comparisons with past findings, the earliest in the list being that Saloot et al. in 2014 (Athirah et al., 2019; Chekima & Alfred, 2017; Saloot et al., 2014). The five top words are "*ni*" (this), "*nk*" (want), "x" (negation), "*nie*" (this), and "*je*" (only). Both "*ni*" and "*nie*" are shortened versions of "*ini*" but Saloot et al. (2014) found only "*ni*", and it ranks third in their corpus. The present study is conducted some years after Saloot et al. (2014), and it is possible that more variations in shortening of "*ini*" has emerged. As for the shortened versions of the verb "*hendak*" (want), "*nak*" ranks higher (number 2) than "*nk*" (number 13) (Saloot et al., 2014). By the time the data were collected for the present study in 2019, only "*nk*" is used (rank of 2) and "*nak*" does not appear in the top 100 neologisms analysed. Athirah et al. (2019) also listed only "*nk*" among the top 10 frequently used words in their analysis of tweets and Facebook comments. This comparison reveals that as time moves on, there is a shift towards using the shorter version "*nk*". Interestingly, "*je*" (shortened version of "*sahaja*") is stable and remains frequently used (rank of 9 in Saloot et al., 2014; rank of 3 in Athirah et al., 2019; rank of 5 in the present study). The alternative spelling "*jer*" was only found by Chekima and Alfred (2017). Since there are indications of changes over time, our corpus of neologisms would be useful to researchers and agencies, like *Dewan Bahasa dan Pustaka*, interested in the evolution of the Malay language in cyberspace.

Secondly, the main morphological process for the formation of neologisms in Malay tweets are clipping and pseudo-elliptical construction, with little use of blending. The economy of time and space in social media communication explains why 98 of the top 100 neologisms are shortened and only two words are elongated. Clipping also accounts for a majority of neologisms in Omar et al.'s (2016) study on Malay tweets but their database is too small for a good comparison (only 15 neologisms). Many researchers have found clipping in their studies on Malay social media discourse (Nisa, 2016; Omar et al., 2016; Paputungan, 2015; Sukma, 2014). A feature that has not been highlighted by other researchers is that clipping of vowels occur more frequently than consonants and syllables. For Malay words involving the consonant-vowel-consonant pattern, consonants are less likely to be removed because a consonant frame is important to make the

truncated words recognisable. The consonant-vowel spelling pattern also prevails in words formed through blending, such as "*tujela*" from "*itusahajalah*" (that's all). The underlined parts of two out of three words ("*itu*" and "*sahaja*") are combined with "*lah*" (without the "h") to form "*tujela*". "*Tujela*" has the consonant-vowel pattern ("*tu*", "*je*", and "*la*"). Although tweets are written, the blended words reflect how the words are spoken. Interestingly, none of the top 100 words in Malay tweets are formed through acronym/initialism and reduplication (e.g., haha, huhu) which is common in other studies (e.g., Kilim Nam & Jung, 2022; Thelwall, 2021).

Thirdly, our study shows that taking account of the phonological aspect can explain the pseudo-elliptical constructions of neologisms in Malay tweets. Researchers like Craig (2003) and Izazi and Tengku-Sepora (2020) have the inanities category for non-classifiable neologisms while Nkhata and Jimaima (2020) has the pseudo-elliptical construction category to group neologisms that flout known word formation processes. The pseudo-elliptical constructions involving letters of the alphabet for negation (X) and the word "*pergi*" (go) are not novel findings ("g" and "p"). However, we explored the possible phonological rules that apply to the output of the word formation, because this aspect has been neglected in other studies on formation of neologisms in social media discourse. We found phonological adjustments that can explain vowel and consonant changes in words classified as having pseudo-elliptical constructions. On vowel changes, we show that there is a simplification of diphthongs to monophthongs, and a movement from vowels in the closed position towards vowels in the middle position [e] or [o]. As for consonant change, our findings show that there is a movement in the place of articulation towards bilabial consonants [p], [m], and [w] but if the original words contain [h] and [r], they are likely to be dropped or replaced. The changes do not affect the voicing quality of the consonants.

CONCLUSION

Our study on youth netspeak in Malay Twitter posts shows that rule-governed nature of lexical innovations underlying neologisms. The top five neologisms were determiners and general verbs, namely, ("*ni*" (this), "*nk*" (want), "x" (negation), "*nie*" (this), and "*je*" (only). The most common lexical morphological processes for the formation of neologisms in social media discourse within the scope of tweets posted by Malaysian Twitter users are clipping and pseudo-elliptical construction, with little use of blending. Using Kiparsky's (1983) theory of lexical morphology and phonology, we have shown that the phonological aspect can explain the pseudo-elliptical constructions of neologisms in Malay tweets.

These phonological adjustment rules are useful to researchers working on the formulation of algorithms to normalise neologisms in Malay social media discourse. By understanding these rules, researchers can design algorithms that are able to accurately recognise, and process Malay neologisms used by social media users, which can then contribute to the enhancement of various applications such as speech recognition, part-of-speech tagger, machine translation, and sentiment analysis. As reiterated by Ariffin et al. (2018), these applications related to the Malay language remains a challenge due to the current confined efforts on standard Malay corpora without considering neologisms as presented in this study. Thus, the findings are able to assist researchers to generate suitable solutions in normalising neologisms in netspeak.

However, a limitation is that the findings are applicable to Malay. Malay has only the consonant-vowel-consonant word structure. As different languages have different spelling systems, we suggest that future studies focus on investigating phonological adjustment rules to find out if addressing the phonology aspect can better explain the formation of neologisms in other

languages. Furthermore, this study highlights the need for a larger corpus involving a greater diversity of users in order to accurately analyse and understand the way in which the language is changing and evolving in various social media platforms, not limited to Twitter. Beyond technological implications, the prevalent use of various lexical innovation processes among the younger generation as presented in this study shows that neologisms can help to maintain the vitality and vibrancy of the Malay language, and ensure that it remains an expressive means of communication in the digital world. Youth in Malaysia are using neologisms and other forms of lexical innovation to express themselves and to keep up with the rapid pace of technological and societal change. Thus, this study underscores the importance of understanding and studying lexical innovation processes in the Malay language, and the role that social media platforms like Twitter play in shaping the evolution of language.

ACKNOWLEDGEMENT

The authors acknowledge the support of Research, Innovation & Enterprise Centre, Universiti Malaysia Sarawak for the research.

REFERENCES

Anthony, L. (2005). AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. *Proceedings of Professional Communication Conference*, pp. 729-737. 10.1109/IPCC.2005.1494244.

Ariffin, S. N. A. N., & Tiun, S. (2018). Part-of-speech tagger for Malay social media texts. *GEMA Online Journal of Language Studies*, *18*(4), 124-142

Athirah, R. R., Soon, L. K., & Haw, S. C. (2019). Exploring edit distance for normalising out-of-vocabulary Malay words on social media. In R. Černý, J. Kočí & V. Kočí (Eds.), *MATEC Web of* Conferences (pp. 1-7). Les Ulis, France: EDP Sciences.

Barseghyan, L. (2013). On some aspects of Internet slang. *Graduate School of Foreign Languages N, Yerevan State University, Yerevan, 14*, 19-31. Retrieved December 20, 2020, from http://ysu.am/files/Liana%20BARSEGHYAN.pdf

Baldwin, T., & Li, Y. (2015). An in-depth analysis of the effect of text normalization in social media. Proceedings of the *2015 conference of the North American chapter of the association for computational linguistics: human language technologies* (pp. 420-429).

Bauer, L. (2006). Compounds and minor word-formation types. In B. Aarts & A. McMahon (Eds.), *The handbook of English linguistics* (pp. 463-482). Blackwell Publishing.

Chekima, K., & Alfred, R. (2017). Sentiment analysis of Malay social media text. In R. Alfred, H. Iida, A. A. A. Ibrahim, & Y. Lim (Eds.), *Proceedings of the Fourth International Conference on Computational Science and Technology* (pp. 205-219). Singapore: Springer.

Craig, D. (2003). Instant messaging: The language of youth literacy. *The Boothe Prize Essays* 1, 116-133. Retrieved December 20, 2020, from https://nl.ijs.si/janes/wp-content/uploads/2014/09/craig03.pdf

Gligorić, K., Anderson, A., & West, R. (2018). *How constraints affect content: The case of Twitter's switch from 140 to 280 characters*. Retrieved December 20, 2020, from https://arxiv.org/abs/1804.02318

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.

Harared, N. (2018). Slang created and used in 1CAK. COM Site: A sociolinguistics study. *Humaniora, 9*(2), 119-127.

Hong, L., Convertino, G., & Chi, E. H. (2011). Language matters in Twitter: A large scale study. In L. A. Adamic, R. A. Baeza-Yates, & S. Counts (Eds.), *Proceedings of the Fifth International Conference on Weblogs and Social Media*. Catalonia, Spain: The AAAI Press.

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), 46-58.

Jaidka, K., Zhou, A., & Lelkes, Y. (2019). Brevity is the soul of Twitter: The constraint affordance and political discussion. *Journal of Communication, 69*(4), 345-372.

Kamus Dewan (4th ed.). 2007. Kuala Lumpur, Malaysia: Dewan Bahasa dan Pustaka.

Kilim Nam, J. A., & Jung, H. Y. (2022). The emergence and spread of Korean COVID-19 neologisms in news articles and user comments and their lexicographic. *Lexicography of Coronavirus-related Neologisms, 163,* 43-72.

Kiparsky, P. (1983). Word-formation and the lexicon. In F. J. Ingemann (Ed.), *Mid-America Linguistics Conference Papers (*pp. 3-29). Dept. of Linguistics, University of Kansas.

Klymenko, O. (2019). Twitterverse: The birth of new words. *Proceedings of the Linguistic Society of America, 4*(11), 1-12.

Ljubotenska, D. (2018). The influence of instant messages on English word formation and changes in the dynamics of the communication process. *Knowledge International Journal, 26*(3), 785-788.

Lutfiyana, F. (2020). *Students' accuracy in pronouncing English palate alveolar sounds: A case study of students of English Department, Universitas Negeri Semarangin the Academic year of 2019/2020*. Final Year Project report, Universitas Negeri Semarangin.

Mustafa, S. Z. B., Kandasamy, M., & Yasin, M. S. M. (2015). An analysis of word formation process in everyday communication on Facebook. *International Journal of Education and Research, 3*(6), 261-274.

Mutiah, S. D., Hidayat, D. N., & Alek, A. (2021). Exploring the word formation process on Rich Brian's official music video comments. *EduLite: Journal of English Education, Literature and Culture, 6*(1), 198-212.

Mworia, R. (2015). *Use of English neologisms on social media: A case of Twitter language in Kenya*. Nairobi: University of Nairobi.

Nevihayati, M. (2018). *English slang word formation processes on twitter*. Doctoral dissertation, Universitas Negeri Malang.

Ningsih, W. W. (2020). *Word formation analysis of slang language used in comments of Rich Brian's Youtube channel.* Final year project report, State Institute for Islamic Studies, Salatiga, Indonesia.

Nisa, R. F. (2016). *Word formation process of non-standard vocabulary in Twitter statuses of Indonesian dangdut singers*. Final Year Project Report, Diponegoro University.

Nkhata, L., & Jimaima, H. (2020). Neologisms: A morphological analysis of social media discourses on the Zambian online media. *Multidisciplinary Journal of Language and Social Sciences Education, 4*, 66-93.

Omar, S. N., Najib, A. F. M., & Zaihan, Z. A. M. (2016). New-word formation processes in Twitter: A study among Malaysian youth. Retrieved January 10, 2021, from https://www.academia.edu/35623710/NEW-

WORD_FORMATION_PROCESSES_IN_TWITTER_A_STUDY_AMONG_MALAYSIAN_YOUTH.docx

Paputungan, A. U. (2015). Analysis of word formation process in Twitter. *Skripsi*, *1*. Retrieved January 10, 2021, from https://repository.ung.ac.id/skripsi/show/321411087/analysis-of-word-formation-process-in-twitter.html

Parkins, R. (2012). Gender and emotional expressiveness: An analysis of prosodic features in emotional expression. *Pragmatics and Intercultural Communication*, *5*(1), 46–54.

Pesetsky, D. (1979). *Russian morphology and lexical theory*. MIT, Cambridge, Massachusetts.

Pesetsky, D. (1985). Morphology and logical form. *Linguistic Inquiry, 16*(2), 193-246.

Renner, V. (2019). French and English lexical blends in contrast. *Languages in Contrast, 19*(1), 27-47.

Renner, V. (2020). An ecosystem view of English word-formation. *The Mental Lexicon, 15*(1), 4-20.

Rusli, N. F. M., Jobar, N. A., & Adam, N. F. M. (2022). Kesalahan ejaan murid berdasarkan bunyi konsonan dalam ujaran guru: Analisis Fonologi Generatif. *Issues in Language Studies*, *11*(2), 37-60.

Saloot, A. M., Idris, N., Aw, A. A., & Thorleuchter, D. (2014). Twitter corpus creation: The case of a Malay Chat-style-text Corpus (MCC). *Digital Scholarship in the Humanities, 31*(2), 227-243.

Scholarly Community Encyclopedia. (2022). *LOL*. https://encyclopedia.pub/entry/28581

Statista. (2021). Twitter's user base in Malaysia amounts to approximately 5.51 million users. Retrieved July 31, 2021, from https://www.statista.com/forecasts/1145758/twitter-users-in-malaysia

Sukma, D. (2014). Analysis of word formation in Short Message Service (SMS) text sent by the second semester students at Pasir Pengaraian University. *Edu Research, 3*(1), 43-52.

Suparta, I. W., Qomariana, Y., & Rahayuni, N. K. S. (2016). Word formation on Twitter posts. *Jurnal Humanis,18*(2), 6-23.

Tchokni, S., Seaghdha, D., & Quercia, D. (2014). Emoticons and phrases: Status symbols in social media. *Eighth international AAAI conference on weblogs and social media*. Palo Alto, CA.

Thelwall, M. (2021). This! Identifying new sentiment slang through orthographic pleonasm online: Yasss slay gorg queen ilysm. *IEEE Intelligent Systems, 36*(4), 114-120.

Thorpe, S. (2002). Vowel-consonant patterns (Part 1). *Word Ways, 35*(4), 249-254.

Twitter. (2020). Retrieved December 20, 2020, from https://www.investopedia.com/terms/m/monthly-active-user-mau.asp

Wu, L., Morstatter, F., & Liu, H. (2018). SlangSD: Building, expanding and using a sentiment dictionary of slang words for short-text sentiment classification. *Language Resources and Evaluation, 52*(3), 839-852.

Yang, J. S., Ko, M. S., & Chung, K. S. (2019). Social emotional opinion decision with newly coined words and emoticon polarity of social networks services. *Future Internet, 11*(8), 165-171.

ABOUT THE AUTHORS

Nur Nashatul Nasuha Nazman graduated with a Bachelor of Arts in Linguistics and a Masters of Applied Linguistics from Universiti Malaysia Sarawak. Her research interest is in word formation and social media communication.

Su-Hie Ting teaches research methodology and is a professor at the Faculty of Language and Communication, Universiti Malaysia Sarawak. She holds a PhD in Applied Linguistics from the University of Queensland. She has published on language choice and identity, academic writing, communication strategies, and health communication. Contact: shting@unimas.my, suhieting@gmail.com

Kee-Man Chuah is a senior lecturer at the Faculty of Language and Communication, Universiti Malaysia Sarawak. He graduated with a Bachelor of Science (Hons) in Teaching of English as a Second Language and a Master of Science in Cognitive Science from Universiti Malaysia Sarawak. His areas of expertise lie in virtual reality, e-learning, and computer-assisted language learning.

eISSN: 2550-2131
ISSN: 1675-8021