

YOUTH SLANG AS SOCIAL, ECONOMIC AND POLITICAL INDEXES IN MALAYSIA AND INDONESIA

Ahmad Syafiq Amir Abdullah Zawawi¹, Hishamudin Isam², N. Lia Marliana³

Nanyang Technological University Singapore, Universiti Utara Malaysia, Universitas Negeri Jakarta

Email: syafiqamir2017@gmail.com , din@uum.edu.my, nliamarliana2929@gmail.com

ABSTRACT

This paper examines youth slang as a sociolinguistic index of social, economic and political concerns in Malaysian and Indonesian digital communication. Drawing on a web-scraped and classified dataset of slang items from Instagram and Threads, the study compares issue patterns across country, platform and gender categories. The analysis applies descriptive frequency reading and qualitative thematic interpretation to three (3) main domains: social interaction, economic experience and political discourse. The findings show that social slang is dominated by conversational response, peer affiliation, emotional self-positioning and digital identity. Economic slang indexes both financial strain and consumer aspiration through terms related to cost of living, online shopping, paylater culture, gig work and budgeting. Political slang is especially marked by criticism, polarization, integrity concerns and campaign narratives. Indonesia records higher overall issue frequencies, while Malaysia shows strong localised forms in cost-of-living and partisan labels. The paper argues that youth slang should not be treated merely as informal vocabulary, but as a compact semiotic resource through which young users interpret everyday pressures, community membership and public life. The study contributes to digital sociolinguistics and language education by foregrounding slang as evidence of contemporary social meaning.

Keywords: *youth slang; digital discourse; Malaysia; Indonesia; sociolinguistics*

INTRODUCTION

Youth slang is often viewed as playful, unstable and peripheral to formal language study. In digital communication, however, slang performs a much more complex function. It enables young users to mark intimacy, evaluate public events, dramatise personal experience, signal group belonging and compress social judgement into short, recognisable expressions. This position is consistent with Crystal (2006) argument that internet communication expands the creative possibilities of language rather than simply weakening linguistic standards. Terms such as *bestie*, *spill*, *healing*, *pokai*, *paylater*, *buzzer*, *walaun* and *toxic politics* are not simply decorative words in online conversation. They operate as small but powerful signs that connect language to everyday life. In this

sense, slang is a valuable entry point for understanding how young people in Malaysia and Indonesia make meaning in networked spaces.

The rise of social media has intensified the visibility of youth language. Social network sites provide technological and interactional affordances that shape how users display relationships, construct profiles and participate in public communication (Boyd & Ellison, 2007). Instagram foregrounds visual self-presentation, lifestyle narration, affective display and consumer culture, while Threads encourages more text-based reactions, commentary, debate and conversational alignment. These platform affordances shape the kinds of slang that become salient. The same broad youth culture may circulate across both spaces, but the pragmatic functions of its slang are not identical. A term such as *vibes* may attach itself to aesthetic identity on Instagram, while *hot take* and *thread panjang* are more closely associated with opinion, argument and public response on Threads. Slang therefore needs to be read not only as a lexical phenomenon but also as platformed discourse.

Malaysia and Indonesia provide a particularly productive comparison. The two countries share historical, linguistic and cultural affinities within the Malay-Indonesian language ecology, yet their digital youth cultures also show national, urban, political and consumer-specific differences. Blommaert (2010) argues that globalisation makes linguistic resources mobile across spaces while also changing their value in local contexts. This insight is useful for understanding why some forms travel easily across borders, especially English-derived items such as *bestie*, *spill*, *healing*, *glow up*, *red flag*, *paylater* and *side hustle*. Other forms are more locally anchored. Malaysian political slang includes expressions such as *walaun*, *walanon*, *macai*, *cai* and *dedak*, while Indonesian data contain forms such as *buzzer*, *cebong*, *kampret*, *kadrun*, *politik toxic* and *capres*. These differences show that shared digital culture does not erase local histories of public debate.

The research problem addressed in this paper is the tendency to describe slang only as evidence of linguistic informality, lexical creativity or language decline. Such descriptions are incomplete because slang also indexes the issues that matter to young users. In the dataset analysed here, the highest-frequency items do not appear randomly. They cluster around social belonging, emotional self-management, online conversation, cost-of-living pressure, consumer desire, small-scale entrepreneurship, political criticism and concerns about integrity. A slang item may be short, humorous or ironic, but its recurrence points to wider patterns in youth experience. This follows Eckert (2012) view that linguistic variation can operate as a social-semiotic system through which speakers produce meaning. The question is therefore not merely what slang words are used, but what social meanings are repeatedly made visible through those words.

The study is guided by three research questions. First, what major social, economic and political issues are indexed by youth slang in Malaysian and Indonesian digital communication? Second, how do the issue patterns differ across the two countries and between Instagram and Threads? Third, what do these patterns suggest about the sociolinguistic role of slang in youth digital discourse? These questions are important because digital slang increasingly enters classrooms, public discourse, advertising, activism and everyday intergenerational interaction. Understanding it systematically can support language education, media literacy and sociolinguistic awareness.

Theoretically, the paper draws from digital sociolinguistics and computer-mediated discourse analysis. Herring (2004) argues that computer-mediated discourse must be studied through observable language use in its social and technological contexts.



Androutsopoulos (2008) similarly emphasises the value of online discourse as situated social practice rather than disembodied text. From an identity perspective, Bucholtz and Hall (2005) conceptualise identity as something produced through semiotic practice, relational positioning and social evaluation. Slang fits this view because it enables speakers and writers to align with certain groups, distance themselves from others and evaluate people or events through condensed labels. It also fits Blommaert (2010) emphasis on the mobility of linguistic resources across scales of social life.

Work on social media further shows that online language is shaped by imagined audiences, context collapse and ambient affiliation. Marwick and Boyd (2011) show that users produce content while negotiating overlapping audiences, while Zappavigna (2011) argues that microblogging allows users to form affiliation through shared evaluative language. The wider social networking environment described by Boyd and Ellison (2007) also helps explain why slang becomes visible as a marker of participation, profile performance and networked sociability. Youth slang has a similar affiliative function. A user who writes *spill*, *valid*, *real* or *no cap* does not only transmit information; the user also takes a stance and invites recognition from others who understand the shared code. Slang therefore works as a mechanism of rapid social alignment.

This paper contributes to these discussions by focusing on issue classification rather than only individual word meaning. The analysis treats slang as a bridge between micro-level linguistic forms and macro-level social concerns. It shows how social, economic and political experiences become linguistically compacted in digital youth vocabulary. By treating recurring slang as indexical evidence, the study extends the sociolinguistic concern with social meaning in variation (Eckert, 2012) into the context of Malay-Indonesian digital discourse. The contribution is especially relevant to Malay and Indonesian language education because it demonstrates that slang can be studied as a structured field of meaning, not merely as informal expression outside legitimate academic attention.

METHOD

This study uses a qualitative descriptive design supported by quantitative frequency comparison. The data were obtained through a web scraping procedure conducted on publicly accessible Instagram and Threads content. Web scraping is understood here as the automated retrieval, parsing and storage of web-based information for research purposes (Munzert et al., 2015; Mitchell, 2024). In digital social research, scraping is useful because it enables researchers to observe naturally occurring online discourse, but such data must be interpreted in relation to platform structure, access conditions and ethical responsibility (Marres & Weltevrede, 2013; Salganik, 2018).

The scraping procedure was designed to identify youth slang appearing in Malaysian and Indonesian digital communication. Public textual content was collected through keyword-based searches and platform-specific browsing paths related to three thematic domains: social interaction, economic experience and political discourse. The collection focused on visible textual materials such as posts, captions, short conversational entries and other public language traces that contained relevant slang forms. Private accounts, direct messages, login-restricted material and personally

identifying information were excluded from the analytical dataset. This decision follows the view that social media data collection must consider not only technical accessibility but also user privacy, platform terms and possible research harm (Brown et al., 2025; Krotov et al., 2020).

The choice of web scraping was also shaped by the current post-API environment of social media research. Earlier studies have noted that application programming interfaces can support structured social media data collection, but API access is increasingly restricted, unstable or uneven across platforms (Lomborg & Bechmann, 2014). The reduction of free or affordable API access has encouraged researchers to consider alternative data collection strategies, including scraping, while maintaining methodological transparency and ethical caution (Freelon, 2018; Trezza, 2023). In this study, web scraping was therefore used not as an uncontrolled harvesting technique, but as a targeted procedure for collecting publicly visible language data relevant to youth slang.

Data collection involved several operational steps. First, a list of slang items and issue-related seed terms was prepared from Malay, Indonesian and English-influenced youth vocabulary. Second, these terms were used to locate public Instagram and Threads content associated with social, economic and political themes. Third, the scraper retrieved only the textual fields needed for linguistic analysis. Fourth, the retrieved data were stored in a structured format to allow subsequent cleaning, counting and classification. The procedure follows the general logic of automated data collection, in which retrieval, parsing, storage, validation and preparation are treated as connected stages rather than isolated technical tasks (Munzert et al., 2015; Mitchell, 2024).

The scraped data were then cleaned before analysis. Exact duplicates, irrelevant advertisements, broken text fragments, repeated spam-like entries and non-substantive material were removed. At the same time, non-standard spelling, code-mixing, abbreviations, reduplication, creative punctuation and platform-specific forms were retained because these features are central to youth slang and digital discourse. This step was important because over-cleaning can erase sociolinguistic evidence. The cleaning process therefore balanced technical standardisation with the need to preserve the informal and indexical character of online language.

After cleaning, the data were organised according to country, platform, thematic domain and gender category as represented in the dataset. The main countries were Malaysia and Indonesia, while the platforms were Instagram and Threads. The three major domains were social, economic and political. Within these domains, the slang items were classified into more specific issue categories, including peer relationship and digital community, conversation and digital discourse, emotional well-being, identity and aesthetics, cost-of-living pressure, online purchasing, gig economy, consumer culture, political polarisation, integrity-related discourse and campaign narratives. The classification follows a discourse-centred approach in which lexical items are interpreted through their social meaning and platform context (Androutsopoulos, 2008; Herring, 2004).

Frequency analysis was used to identify dominant slang items and issue categories. The dataset records the 15 highest slang items by country, platform, theme and gender category, with each item accompanied by frequency values, percentages and issue classification. These values were used descriptively, not inferentially. In other words, frequency was treated as an indicator of prominence within the dataset, not as proof of universal usage among all Malaysian and Indonesian youth. This is important



because web-scraped data may be affected by platform visibility, search pathways, algorithmic ranking, data availability and the temporal conditions of collection (Brown et al., 2025; Marres & Weltevrede, 2013).

The qualitative stage involved interpretive reading of high-frequency slang items across the selected categories. The analysis examined how slang indexes social affiliation, emotional self-positioning, consumer desire, financial strain, gig work, political judgement and public distrust. Instagram and Threads were treated as different communicative environments rather than neutral containers of data. Instagram was interpreted as a more visual and identity-oriented platform, while Threads was interpreted as a more text-driven environment for commentary, reaction and debate. This platform-sensitive interpretation is consistent with social media research methods that emphasise the relationship between data, interface, communicative practice and context (Sloan & Quan-Haase, 2017; Trezza, 2023).

Ethical care was applied throughout the process. The analysis used aggregated lexical and issue-level patterns rather than individual user profiles. Usernames, profile identifiers, URLs and other unnecessary personal traces were not reported. Examples were discussed at the level of slang forms and thematic categories rather than through identifiable user statements. The study therefore follows responsible big data and web scraping principles by reducing privacy risk, limiting the scope of collection, respecting the public-private ambiguity of online communication and avoiding claims that exceed the dataset (Brown et al., 2025; Krotov et al., 2020; Salganik, 2018).

The study has several methodological limitations. First, web scraping captures accessible digital traces rather than the full population of youth language use. Second, social media platforms are dynamic environments; search results, interface design, ranking systems and public availability may change over time. Third, some slang items are polysemous and may shift meaning across contexts, platforms and communities. To address these limitations, the study combines frequency-based description with qualitative thematic interpretation, and all findings are presented as patterns within the collected dataset rather than as exhaustive representations of Malaysian and Indonesian youth speech.

RESULTS AND DISCUSSION

The consolidated dataset contains 1,005,042 issue-level occurrences across the three major domains. Social issues account for 417,586 occurrences, or 41.5 percent of the total. Political issues account for 299,342 occurrences, or 29.8 percent, while economic issues account for 288,114 occurrences, or 28.7 percent. Indonesia records a higher overall frequency than Malaysia across the three domains, with 534,559 occurrences compared to Malaysia's 470,483. This difference does not simply mean that Indonesian youth use more slang in all situations. Rather, it suggests that the Indonesian subset in this dataset shows stronger concentration in the selected high-frequency terms, especially in emotional social vocabulary, gig-economy language and political polarisation.

Table 1. Overall distribution of issue-level slang occurrences by country and theme

Theme	Malaysia (J)	Indonesia (J)	Total (J)	Share of all themes (%)
Social	201,585	216,001	417,586	41.5
Economic	132,234	155,880	288,114	28.7
Political	136,664	162,678	299,342	29.8
Total	470,483	534,559	1,005,042	100.0

Table 1 shows that social slang is the largest domain in the dataset. This is expected because youth digital communication is fundamentally relational. Young users do not merely use social media to publish information; they use it to greet, react, joke, evaluate, affirm, tease, narrate and seek recognition. The social domain therefore includes both interactional markers such as *weh*, *bro*, *sis*, *bestie*, *ngab* and *cuy*, and stance markers such as *spill*, *hot take*, *valid*, *real* and *no cap*. These items enable users to create a shared conversational rhythm. Their value lies not in literal reference alone but in the ability to signal membership in a recognisable digital style.

At the same time, the political and economic domains are almost equally prominent. This is important because it complicates the common assumption that youth slang is mostly playful or entertainment-oriented. Political slang accounts for nearly one third of the total, while economic slang accounts for slightly less than one third. Together, these two domains show that young users use informal vocabulary to process serious public and material concerns. Slang offers a way to speak about economic pressure, political disappointment and social distrust in forms that are emotionally immediate and shareable.

Table 2. Dominant issue categories within social, economic and political themes

Theme	Dominant issue category	Malaysia (J)	Indonesia (J)	Total (J)	Share within theme (%)
Social	Conversation, response and digital discourse	51,020	50,870	101,890	24.4
Social	Peer relationship and digital community	43,607	31,331	74,938	17.9
Social	Emotion, mental well-being and self-relationship	21,476	46,098	67,574	16.2
Economic	Cost of living and financial pressure	52,086	33,875	85,961	29.8
Economic	Consumer lifestyle and spending culture	29,720	34,726	64,446	22.4
Economic	Side income, gig economy and small business	12,178	40,754	52,932	18.4
Political	Criticism, activism and political polarisation	53,104	62,259	115,363	38.5
Political	Integrity, corruption and political manipulation	41,028	39,187	80,215	26.8
Political	Campaign narratives and digital debate	22,404	34,409	56,813	19.0

SOCIAL SLANG: AFFILIATION, AFFECT AND PLATFORMED CONVERSATION

The social theme is dominated by conversation, response and digital discourse, with 101,890 occurrences or 24.4 percent of all social issue classifications. This category includes items such as *hot take*, *thread panjang*, *utas panjang*, *spill*, *deep talk*, *valid*, *ratio*, *real*, *no cap* and *random thought*. These forms show that youth slang is strongly tied to the management of participation. A user does not only post a statement; the user also



frames the statement as an opinion, a confession, a reaction, a story, a provocation or a claim to truth. Such evaluative framing supports Zappavigna's (2011) account of ambient affiliation, where users build community through shared stance. Slang therefore helps users organise the flow of digital talk.

The category of peer relationship and digital community is the second largest social issue, with 74,938 occurrences. Malaysian Instagram data include *weh*, *bro*, *sis*, *bestie* and *masuk geng*, while Indonesian Instagram data include *ngab*, *cuy* and *bestie*. These words serve as address terms and solidarity markers. Their function resembles what sociolinguistic theory describes as identity work: they place users in relation to others and help construct similarity, closeness and informality (Bucholtz & Hall, 2005). The use of *bestie* across both countries is especially significant because it shows the circulation of global English-based intimacy within Malay-Indonesian digital spaces, a process that also reflects the mobility of language under globalisation (Blommaert, 2010).

Emotional and self-related slang is also highly visible, especially in Indonesia. The category records 67,574 occurrences overall, but Indonesia contributes 46,098 compared with Malaysia's 21,476. This pattern is reflected in items such as *healing*, *insecure*, *overthinking*, *salting*, *baper*, *delulu*, *trust issue* and *self love*. These words suggest that youth digital discourse does not separate language from mental and emotional life. Instead, it turns feelings into recognisable social tokens. A term such as *healing* may refer to recovery, leisure, escape or self-care, depending on context. Its recurrence shows how therapeutic vocabulary has become normalised in informal youth expression. In this sense, affective slang can be read as a form of stance and self-positioning within networked audiences (Marwick & Boyd, 2011).

Instagram and Threads show different social orientations. Instagram contains more vocabulary of image, aesthetics and persona, for example *slay*, *vibes*, *glow up*, *photodump*, *aesthetic*, *outfit check*, *soft girl*, *gemoy* and *main character*. These items are linked to visual identity and the performance of a curated self. Danesi's (2017) discussion of visual language in internet communication is useful here because many identity-related slang items are embedded in image captions, emoji use and aesthetic display. Threads, in contrast, contains more vocabulary of textual engagement, for example *hot take*, *ratio*, *thread panjang*, *utas panjang*, *bait*, *reply guy*, *valid* and *deep talk*. This difference supports the view that platform affordances shape discourse practices (Boyd & Ellison, 2007). The platform does not determine language mechanically, but it encourages certain kinds of participation and therefore certain kinds of slang.

The meme and humour category, including *cringe*, *flex*, *anjay*, *skena*, *red flag* and *takleh brain*, shows another important social function. These terms allow users to evaluate behaviour indirectly. Calling something *cringe* or *red flag* is a compact act of social judgement. It may be humorous, but it also establishes norms about what is embarrassing, excessive, suspicious or unacceptable. Meme-oriented slang therefore works as a distributed form of moral commentary, where young users collectively assess behaviour through playful labels. This interpretation is supported by Shifman (2014) view that internet memes circulate through imitation, transformation and shared cultural recognition, and by Tagg et al. (2017), who show how online interaction can generate offence, alignment and disagreement through everyday evaluative practices.

The social findings indicate that youth slang should be understood as an ecology of affiliation and evaluation. It helps users enter conversations, maintain closeness, display affect, present identity and judge social behaviour. These functions are consistent with the view that linguistic variation produces social meaning rather than merely reflecting fixed social categories (Eckert, 2012). In educational contexts, this means that slang can be used as material for discussing pragmatics, discourse community, semantic change and digital citizenship. Rather than simply correcting slang, educators can guide students to analyse when, why and how informal forms carry social meaning.

ECONOMIC SLANG: FINANCIAL STRAIN, CONSUMER DESIRE AND DIGITAL WORK

The economic theme reveals a strong connection between slang and material life. The largest economic category is cost of living and financial pressure, with 85,961 occurrences or 29.8 percent of economic issue classifications. Malaysian data are especially strong in this category, with 52,086 occurrences compared to Indonesia's 33,875. Items such as pokai, sengkek, auto pokai, kopak, duit tepi, harga kayangan, cekik darah and auto miskin show that financial strain is made visible through humorous and exaggerated expressions. These words allow young users to talk about economic difficulty without necessarily adopting formal economic vocabulary. Following Crystal (2006), such forms demonstrate the creative adaptation of language to new communicative environments.

The Malaysian cost-of-living lexicon is particularly expressive. Pokai and sengkek index being short of money, while kayangan and harga kayangan dramatise unaffordable prices. Cekik darah evaluates a price as exploitative or unreasonable. Such terms are not neutral descriptions; they contain affective judgement. They show how slang compresses an economic complaint into a socially recognisable phrase. The frequency of these terms suggests that young Malaysians use slang to negotiate the tension between limited resources, lifestyle aspiration and everyday expenditure.

Indonesian data also contain strong financial-pressure terms, including bokek, dompet tipis, tanggal tua, gaji numpang lewat, tekor and dompet nangis. These items often use metaphor and humour. Dompet tipis personifies the wallet as thin, while gaji numpang lewat imagines salary as something that merely passes through without staying. The humour does not reduce the seriousness of the issue. Instead, it makes economic vulnerability communicable in a socially shareable way. The use of playful metaphor is therefore a strategy for narrating precarity.

The second largest economic category is consumer lifestyle and spending culture, with 64,446 occurrences or 22.4 percent. This category includes racun belanja, racun shopping, wishlist, haul, thrift, preloved, shop haul and jastip. These terms reveal the close relationship between youth slang and platform commerce. Racun belanja, literally associated with being 'poisoned' into buying, is particularly meaningful because it frames consumption as contagious desire. Wishlist and haul connect self-presentation to purchasing, while thrift and preloved show that consumer culture also includes budget-conscious, sustainable or second-hand practices. In social network environments, consumer terms become part of profile work and peer influence, not merely private purchasing decisions (Boyd & Ellison, 2007; Marwick & Boyd, 2011).

The third economic category, side income, gig economy and small business, records 52,932 occurrences or 18.4 percent. Indonesia is much stronger in this category, with 40,754 occurrences compared to Malaysia's 12,178. Items such as cuan, cuan thread,



modal receh, reseller, affiliate, side hustle, small biz and freelance indicate that youth economic language is not only about shortage. It is also about opportunity, entrepreneurship and monetisation. The digital economy has given young users new ways to talk about income generation, even when these opportunities remain uncertain or small-scale.

The rise of paylater, cashback, free shipping, checkout, COD, promo, diskon and gratis ongkir shows that economic slang is also shaped by e-commerce infrastructure. These terms are not merely borrowed from marketing. They are integrated into youth discourse as part of decision-making, self-justification and social persuasion. Paylater is especially important because it links consumption to deferred payment and normalises small debt within everyday digital purchasing. Its presence in both Malaysia and Indonesia suggests that youth slang is responsive to the financial technologies embedded in platform economies. Such uptake illustrates how linguistic resources travel through global digital systems but acquire situated local meanings (Blommaert, 2010).

The economic findings show that slang functions as a vernacular economy of experience. Formal economics may discuss inflation, purchasing power, labour precarity and digital transactions, but youth slang translates these concepts into everyday affective language. This has implications for consumer education and financial literacy. Terms such as paylater, bajet lari, budgeting, cashflow and tip jimat can become entry points for discussing spending behaviour, debt, budgeting and platform persuasion in ways that connect directly to youth experience.

POLITICAL SLANG: POLARISATION, INTEGRITY AND PUBLIC JUDGEMENT

Political slang is the second largest domain in the overall dataset and the most concentrated in a single issue category. Criticism, activism and political polarisation record 115,363 occurrences, representing 38.5 percent of political issue classifications. This dominance shows that political slang among youth is strongly evaluative. It is used to name, mock, delegitimise, defend and attack political actors or their supporters. Such language may be humorous, but it also participates in the construction of political camps. In this respect, political slang resembles the online stance-taking and offence-management practices discussed by Tagg et al. (2017).

Malaysian political slang is marked by terms such as walaun, walanon, macai, cai, dedak, cybertrooper, keyboard warrior, kaki spin, katak politik, bossku, trooper, kencing manifesto, sembang kencing, kartel and tentera bawang. These items show several layers of political meaning. Some label supporters or partisan identities. Others accuse actors of manipulation, opportunism or propaganda. Dedak, for example, implies being rewarded or fed for political loyalty, while kaki spin evaluates someone as manipulating narratives. Kencing manifesto frames campaign promises as deceptive. These words are not neutral. They are weapons of evaluation within digital political conflict.

Indonesian political slang has a different but comparable structure. Items such as buzzer, cebong, kampret, kadrun, politik toxic, oligarki, dinasti, korupsi, capres, caleg, rakyat kecil, oposisi, boikot and thread fakta point to polarisation, campaign discourse, integrity concerns and social justice. The term buzzer is especially salient because it indexes the perception of organised political amplification or manipulation. Cebong,

kampret and kadrin are factional labels that have circulated in Indonesian political debate, showing how slang can harden group boundaries.

The second largest political category is integrity, corruption and political manipulation, with 80,215 occurrences or 26.8 percent. This category is important because it shows that youth political discourse is not limited to partisan mockery. It also contains recurring concern with trust, corruption, propaganda, cronies, dynastic politics and manipulation. The slang forms may be informal, but the issues they index are central to democratic life. When young users repeatedly use terms related to cronies, cartels, corruption, buzzers and cybertroopers, they are articulating suspicion toward the fairness of political communication and governance. These labels function as indexical signs because they connect a lexical item with a wider field of social judgement (Bucholtz & Hall, 2005; Eckert, 2012).

Campaign narratives and digital debate record 56,813 occurrences or 19.0 percent. This category includes terms such as manifesto, kempen digital, hot take politik, debate, naratif, thread fakta, kaki spin and sembang kencang. Threads is particularly relevant here because the platform supports longer text, replies and public argument. Political slang on Threads often appears as commentary on claims, counterclaims and perceived manipulation. This supports the argument that the discourse environment shapes the kind of political slang that becomes visible.

Political slang also demonstrates the ambivalence of digital youth participation. On one hand, it can signal democratic engagement, critical awareness and willingness to evaluate political actors. On the other hand, it may intensify polarisation by reducing complex political positions to hostile labels. This ambivalence is consistent with wider work on social media publics, where affiliation and conflict are produced through rapid, shareable expressions (Tagg et al., 2017; Zappavigna, 2011). For language education, the implication is that political slang should be analysed as a form of stance-taking that requires ethical and critical literacy.

Table 3. Selected high-frequency slang patterns and sociolinguistic interpretations

Domain	Malaysia examples	Indonesia examples	Interpretive pattern
Peer affiliation	weh, bro, sis, bestie, masuk geng	ngab, cuy, bestie, ngabers	Address terms create closeness and signal digital community membership.
Digital conversation	spill, hot take, thread panjang, valid, no cap	spill, hot take, utas panjang, no debat	Slang frames opinion, disclosure, reaction and public participation.
Emotion and self	healing, insecure, delulu, overthinking	healing, salting, baper, insecure, overthinking	Affective slang makes mental and relational states socially visible.
Cost of living	pokai, sengkek, kayangan, cekik darah	bokek, dompet tipis, tanggal tua, tekor	Financial pressure is expressed through humour, metaphor and exaggeration.
Consumer culture	racun belanja, wishlist, thrift, preloved	racun belanja, promo, diskon, haul, jastip	Platform commerce shapes desire, budgeting and identity performance.
Digital work	duit tepi, side hustle, affiliate, small biz	cuan, reseller, modal receh, affiliate, freelance	Youth economic imagination includes small-scale entrepreneurship and monetisation.
Political polarisation	walaun, walanon, macai, cai	cebong, kampret, kadrin, politik toxic	Labels create political camps and simplify ideological conflict.



Integrity and manipulation	dedak, cybertrooper, kartel, kaki spin	buzzer, oligarki, dinasti, korup	Slang indexes distrust toward propaganda, corruption and elite control.
Campaign discourse	kencing manifesto, sembang kencing, tentera bawang	capres, caleg, thread fakta, janji manis	Digital debate turns campaign claims into evaluative and shareable language.

CROSS-COUNTRY AND PLATFORM COMPARISON

The comparison between Malaysia and Indonesia reveals both convergence and divergence. Convergence is visible in the circulation of transnational youth slang. Items such as *bestie*, *spill*, *healing*, *vibes*, *insecure*, *glow up*, *red flag*, *wishlist*, *thrift*, *cashback*, *paylater*, *affiliate* and *side hustle* show that English-derived and platform-mediated vocabulary travels across national boundaries. These shared forms indicate that youth digital culture in both countries participates in broader global circuits of social media language. This movement of linguistic resources across borders is consistent with Blommaert (2010) account of language mobility in globalisation and Crystal (2006) account of internet-enabled linguistic creativity.

Divergence appears in the way shared domains are locally weighted. Malaysia is particularly strong in cost-of-living slang, while Indonesia is stronger in emotional social vocabulary, gig-economy vocabulary and political polarisation. Malaysian economic items often stress lack, unaffordability and price pressure. Indonesian economic items more frequently combine financial strain with income generation and consumer-platform terminology. These differences may reflect not only national economic conditions but also the way economic anxieties are culturally verbalised in each digital community.

Political divergence is even more pronounced. The Malaysian political lexicon contains highly localised partisan labels and accusations associated with national political culture. Indonesian political slang contains equally localised factional labels, but also a strong presence of terms associated with buzzers, oligarchy, dynastic politics and national electoral vocabulary. The absence of many shared political slang items across the two countries indicates that political vocabulary is more nationally anchored than social or economic vocabulary. Social and consumer slang can travel easily; political slang tends to remain tied to local histories, actors and conflicts.

The platform comparison shows that Instagram is more strongly linked to visual identity, consumer desire and lifestyle expression, while Threads is more strongly linked to textual opinion, commentary and political discussion. This finding aligns with the idea that digital language is shaped by affordances. Instagram encourages the captioning of images, the display of aesthetic selfhood and the circulation of shopping desire. Threads encourages reaction chains, argumentative framing and conversational debate. Youth slang adapts to these conditions by becoming either more image-oriented, more affective, more consumer-oriented or more argumentative depending on the platform. This supports Herring's (2004) and Androutsopoulos's (2008) argument that online language should be analysed in relation to the social and technological conditions of its use.

Gender categories in the dataset also show meaningful but cautious patterns. Female-labelled subsets often contain more items associated with intimacy, emotional language, visual identity and consumer culture, such as *bestie*, *sis*, *glow up*, *cute gila*, *outfit check*, *wishlist*, *racun belanja*, *cashback* and *budgeting*. Male-labelled subsets more

often contain items associated with peer address, lifestyle action, financial struggle, gig income and political antagonism, such as *bro*, *lepak*, *gas*, *pokai*, *cuan*, *macai* and *buzzer*. These patterns should not be read as fixed gender differences. They are better understood as tendencies in how the dataset organises digital participation and self-presentation.

Overall, the cross-country and platform findings support three arguments. First, youth slang is a register of social life because it records peer relations, emotional vocabulary and identity performance. Second, youth slang is a register of economic life because it records cost pressure, consumer desire, platform commerce and digital income practices. Third, youth slang is a register of political life because it records polarisation, distrust, campaign narratives and public judgement. These registers overlap. A term can be humorous and political, emotional and consumerist, or social and economic at the same time.

The educational value of these findings is substantial. In Malay and Indonesian language education, slang is often treated as material to be avoided in formal writing. While this distinction remains important, the present findings suggest that slang also offers authentic material for critical linguistic analysis. Students can examine how meaning shifts across context, how borrowed words become localised, how metaphor expresses social pressure and how digital platforms shape discourse. Such analysis can strengthen awareness of register, audience, ethics and language variation. It also allows students to see, following Eckert (2012), that variation is a resource for producing social meaning rather than merely a deviation from standard usage.

IMPLICATIONS FOR LANGUAGE EDUCATION AND DIGITAL LITERACY

The findings have direct implications for language education. In many classrooms, slang is discussed mainly as vocabulary that should be avoided in formal writing. That advice remains useful for teaching academic register, but it is not enough for contemporary digital literacy. Students also need to understand why informal words become persuasive, how they build solidarity and how they can produce exclusion, stereotyping or conflict. A classroom discussion of terms such as *bestie*, *spill*, *healing*, *paylater*, *buzzer* or *walaun* can become a discussion about audience, stance, context and ethics. This pedagogical move is consistent with computer-mediated discourse analysis, which treats digital texts as socially situated language use (Herring, 2004).

For sociolinguistic teaching, the dataset provides concrete material for demonstrating that language variation is systematic. Students can compare how the same or similar meanings are expressed through different national forms, for example *pokai* and *bokek*, *thread panjang* and *utas panjang*, or *cybertrooper* and *buzzer*. Such comparison helps students see that variation is not linguistic error. It is a pattern of social meaning shaped by local history, media practice and community norms. This is especially valuable in Malay and Indonesian language studies because it links formal linguistic concepts to familiar digital examples.

For semantic and pragmatic teaching, youth slang offers rich evidence of metaphor, evaluation and stance. Economic terms such as *dompet tipis*, *gaji numpang lewat*, *cekik darah* and *harga kayangan* show how metaphor makes economic pressure vivid. Political terms such as *dedak*, *kaki spin* and *politik toxic* show how evaluation is compressed into short labels. Social terms such as *red flag*, *cringe* and *valid* show how users judge behaviour and align with others. These examples can be used to teach connotation, indexical meaning, implicature and semantic change in accessible ways. The inclusion of visual and meme-related forms also allows teachers to discuss emoji, image-



text relations and internet memes as part of contemporary semiotic practice (Danesi, 2017; Shifman, 2014).

For media literacy, the political findings are particularly important. Slang labels can encourage participation, but they can also reduce complex issues to factional insults. When students understand how labels such as macai, cebong, kadrun or keyboard warrior function, they can become more aware of polarising language and manipulation. This awareness does not require banning such terms from discussion. Instead, it requires asking what a label does, whose perspective it represents, what evidence it hides and how it shapes readers' reactions. Such critical questioning is aligned with work on online offence, conviviality and political communication in social media spaces (Tagg et al., 2017).

Finally, the findings suggest that language educators and researchers should continue documenting youth slang as part of living language change. Digital slang changes quickly, but its speed should not be used as a reason to ignore it. On the contrary, rapid change makes it an important site for observing how young users respond to new platforms, new economic pressures and new political events. A critical but open approach can help bridge formal language learning with the realities of everyday digital communication. Future classroom tasks may ask students to classify slang by function, compare meanings across platforms, rewrite informal posts into formal prose and reflect on how register choices affect credibility. Such activities would treat students as analysts of language rather than passive receivers of prescriptive rules, an approach supported by discourse-centred online ethnography and digital discourse analysis (Androutsopoulos, 2008; Herring, 2004).

CONCLUSION

This paper has examined youth slang as a sociolinguistic index of social, economic and political concerns in Malaysian and Indonesian digital communication. The findings show that slang is not merely informal vocabulary or youth entertainment. It is a compact semiotic resource through which young users build affiliation, express emotion, perform identity, evaluate prices, narrate financial pressure, imagine income opportunities, criticise political actors and participate in public debate. These functions confirm that online youth language should be read as socially situated discourse rather than isolated lexical novelty (Androutsopoulos, 2008; Herring, 2004).

The Malaysian and Indonesian comparison shows shared participation in global youth and platform vocabulary, but also strong local anchoring. Social and consumer slang travel across borders more easily, while political slang remains closely tied to national political histories and conflicts. Platform differences are also important. Instagram supports more visual, aesthetic and consumer-oriented slang, whereas Threads supports more discursive, argumentative and political forms. Gendered patterns appear in the dataset, but they should be interpreted as discursive tendencies rather than fixed linguistic traits. Overall, the findings reinforce the importance of analysing identity, audience and affiliation in networked language practices (Bucholtz & Hall, 2005; Marwick & Boyd, 2011; Zappavigna, 2011).

The study contributes to digital sociolinguistics by showing how issue classification can connect lexical forms to broader structures of youth experience. It also contributes to language education by suggesting that slang can be used productively as a resource for teaching register, pragmatics, semantic change, media literacy and critical discourse awareness. Future research should analyse raw posts and interactional sequences to examine how the same slang items function in context, how meanings shift over time and how users themselves explain their choices. Such work would deepen understanding of youth language as a dynamic record of contemporary digital life and extend the study of linguistic variation as social meaning (Eckert, 2012).

ACKNOWLEDGMENT

This section contains thanks to sponsors or fund donors, or to parties who played an important role in carrying out the research.

BIBLIOGRAPHY

- Androutsopoulos, J. (2008). Potentials and limitations of discourse-centred online ethnography. *Language@Internet*, 5.
<https://scholarworks.iu.edu/journals/index.php/li/article/view/37571>
- Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge University Press.
- Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230.
<https://doi.org/10.1111/j.1083-6101.2007.00393.x>
- Brown, M. A., Gruen, A., Maldoff, G., Messing, S., Sanderson, Z., & Zimmer, M. (2025). Web scraping for research: Legal, ethical, institutional, and scientific considerations. *Big Data & Society*, 12(4).
<https://doi.org/10.1177/20539517251381686>
- Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4-5), 585-614.
<https://doi.org/10.1177/1461445605054407>
- Crystal, D. (2006). *Language and the Internet* (2nd ed.). Cambridge University Press.
- Danesi, M. (2017). *The semiotics of emoji: The rise of visual language in the age of the Internet*. Bloomsbury Academic.
- Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, 41, 87-100.
<https://doi.org/10.1146/annurev-anthro-092611-145828>
- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4), 665-668.
<https://doi.org/10.1080/10584609.2018.1477506>



- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338-376). Cambridge University Press. <https://doi.org/10.1017/CBO9780511805080.016>
- Krotov, V., Johnson, L., & Silva, L. (2020). Tutorial: Legality and ethics of web scraping. *Communications of the Association for Information Systems*, 47(22), 539-563. <https://doi.org/10.17705/1CAIS.04724>
- Lomborg, S., & Bechmann, A. (2014). Using APIs for data collection on social media. *The Information Society*, 30(4), 256-265. <https://doi.org/10.1080/01972243.2014.915276>
- Marres, N., & Weltevrede, E. (2013). Scraping the social? Issues in live social research. *Journal of Cultural Economy*, 6(3), 313-335. <https://doi.org/10.1080/17530350.2013.772070>
- Marwick, A. E., & Boyd, D. M. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114-133. <https://doi.org/10.1177/1461444810365313>
- Mitchell, R. (2024). *Web scraping with Python: Data extraction from the modern web* (3rd ed.). O'Reilly Media.
- Munzert, S., Rubba, C., Meissner, P., & Nyhuis, D. (2015). *Automated data collection with R: A practical guide to web scraping and text mining*. John Wiley & Sons.
- Salganik, M. J. (2018). *Bit by bit: Social research in the digital age*. Princeton University Press.
- Shifman, L. (2014). *Memes in digital culture*. MIT Press.
- Sloan, L., & Quan-Haase, A. (Eds.). (2017). *The SAGE handbook of social media research methods*. SAGE Publications.
- Tagg, C., Seargeant, P., & Brown, A. A. (2017). *Taking offence on social media: Conviviality and communication on Facebook*. Palgrave Macmillan.
- Trezza, D. (2023). To scrape or not to scrape, this is dilemma. The post-API scenario and implications on digital research. *Frontiers in Sociology*, 8, Article 1145038. <https://doi.org/10.3389/fsoc.2023.1145038>
- Zappavigna, M. (2011). Ambient affiliation: A linguistic perspective on Twitter. *New Media & Society*, 13(5), 788-806. <https://doi.org/10.1177/1461444810385097>