From Borrowing to Belonging: The Structural Features and Cultural Dynamics of Lexical Innovation in Manglish

Hui Geng [1*], Shameem Rafik Galea [2], Qu Tang [1]

[1*] College of Foreign Studies, Guilin University of Technology, Guilin, China.

[2] Faculty of Education, Languages, Psychology & Music, SEGi University, Selangor, Malaysia.

*Corresponding Author: Hui Geng (Email: huiyuqiankun@gmail.com)

Corresponding Address: College of Foreign Studies, Guilin University of Technology, 12 Jiangan Road, Guilin 541004, Guangxi Zhuang Autonomous Region, China.

Abstract

Manglish, the Malaysian variety of English, embodies the linguistic hybridity emerging from continuous contact among Malay, Chinese, Tamil, and English. While frequently labelled as substandard or deficient, Manglish reveals systematic lexical creativity that reflects Malaysia's multilingual ecology and evolving cultural identity. Despite its sociolinguistic significance, limited empirical work has examined the structural features of lexical innovation in Manglish. This study investigates the types, frequencies, and underlying cultural dynamics of such innovations to uncover how English is structurally and symbolically localized in Malaysia. A 20,000-word corpus was constructed from a classical Manglish book. Texts were cleaned, tokenized, and analyzed through AI-assisted lexical screening (ChatGPT 5.2) and comprehensive manual coding to identify and classify all instances of lexical innovation, based on an integrated analytical framework combining Kalukar et al. (2023) for internal word-formation, Tan (2009) for borrowing and loan translation, and Buschfeld & Kautzsch (2017) for hybridization and localization. Frequencies and proportional distributions were calculated to reveal dominant structural patterns. Findings show that lexical innovation in Manglish is dominated by Borrowing and Loan Translation processes (45.8%), followed by Hybridization and Localization (33.1%), and Word-Formation processes (21.1%). These structural patterns illustrate how English is re-molded through both morphological productivity and contact-induced creativity. The study concludes that Manglish's lexical innovations are adaptive, meaning-laden expressions of identity rather than deviations from Standard English. By bridging structure and culture, the analysis reframes Manglish as a recognized product of linguistic belonging within Malaysia's multilingual context and contributes to broader discussions of creativity

and ownership in World Englishes.

1. Introduction

English never stays still. As it travels across the globe, it continuously evolves, interacting with local languages and cultures. This evolution gives rise to new varieties of English that carry distinct local characteristics and identities. In multilingual societies, English does not replace local languages but intertwines with them, reflecting the unique linguistic dynamics of each society. Malaysia, with its rich mosaic of languages (Malay, Chinese, Tamil, and indigenous languages), provides a compelling example of how English adapts to local conditions, creating a dynamic hybrid language: Manglish. This transformation reflects how people claim ownership of English, reshaping it to fit their needs and experiences. As Schneider (2003) notes, "present-day English as a global language is more than the world's predominant lingua franca—it is also a language which is currently growing roots in a great many countries and communities around the world, being appropriated by local speakers" (p. 233). The evolution of English in Malaysia showcases how language serves as a vehicle for local identity.

Malaysia is a living laboratory of languages, where everyday life is a multilingual performance. From coffee shops to classrooms and markets, Malaysians effortlessly switch between languages, creating a unique linguistic space where words, tones, and cultures collide. Over time, this constant code-switching has produced Manglish, a variety of English that is expressive, humorous, and distinctly local. Rather than a deviation from "standard" English, Manglish is a reflection of Malaysia's ability to blend diverse linguistic and cultural worlds. It is not "broken" English; it is English reshaped for Malaysian life, embodying the country's multicultural and multilingual identity.

Manglish has been the subject of several studies, particularly focusing on its lexical innovation and the pragmatic functions of its discourse markers. Research has identified three key characteristics of Manglish: (i) significant lexical borrowing from Malay, Chinese, and Tamil; (ii) calquing or loan translation of multi-word expressions; and (iii) the frequent use of discourse markers such as *lah*, *lor*, and *meh*, which serve important pragmatic functions like signaling solidarity, formality, and social alignment (Tan, 2009; Baskaran, 2005; Hashim & Tan, 2012). These features are not imperfections but integral parts of a cohesive linguistic system that reflects Malaysia's multicultural landscape. The study of Manglish provides valuable insight into how Malaysians adapt English to fit their local cultural context, offering an understanding of how language evolves in response to social needs.

Despite substantial research on individual features of Manglish, especially lexical borrowing and discourse particles (Tan, 2009; Hashim & Tan, 2012), the field still lacks an integrated structural taxonomy that maps how different innovation mechanisms relate to one another and form a coherent system. Much of the existing research isolates one feature at a time,

leaving underexplored the ways these features co-occur, cluster, and mutually reinforce each other in actual usage. Consequently, no comprehensive framework currently captures how these processes collectively shape Manglish as a distinct variety of English. We understand the individual components, but the pattern that ties them together, and the cultural logic behind that pattern, remains hidden.

This study aims to fill this gap by providing an integrated framework for understanding the processes of lexical innovation in Manglish. By synthesizing key features such as internal morphological processes, contact-induced linguistic extensions, and structural integration, this research will reveal how these elements interact to form a cohesive system that reflects Malaysia's cultural identity. The goal is to provide a deeper understanding of how Manglish evolves through contact with local languages, driven by both linguistic creativity and sociocultural motivations.

2. Literature Review
2.1 The Development and Socio-Cultural Role of Manglish in Malaysia

Manglish, often regarded as a linguistic hybrid, has emerged from centuries of extensive contact between English and Malaysia's primary languages, including Malay, Chinese, Tamil, and various indigenous languages. This variety of English operates on multiple linguistic levels, including lexical, grammatical, and pragmatic levels, and is deeply shaped by Malaysia's socio-political history and multicultural society. As a product of language contact, Manglish has developed into a dynamic, flexible linguistic system that incorporates features from multiple languages, which has led to its characterization as an example of multilingual hybridity (Blommaert, 2010). Pakir (2009) emphasizes that this hybridity reflects the diverse social and cultural realities of Malaysia, where different linguistic traditions have coexisted for centuries, enriching the English language in the process. While it is frequently dismissed as "non-standard" or "substandard" English (Kirkpatrick, 2007), Manglish is an authentic linguistic variety that reflects the cultural and social realities of Malaysia's diverse population. Manglish represents a form of linguistic creativity, wherein speakers adapt English by integrating features from local languages, thereby creating a new, adaptive variety that is inherently tied to local identity and community (Govindan & Pillai, 2009).

As a key vehicle for expressing local identity, Manglish thrives in informal, everyday contexts and is central to the communicative practices of Malaysians, particularly in urban areas, where speakers often switch fluidly between English and local languages. More broadly, Malaysian English can be said to encompass all the sub-varieties of Englishes spoken by Malaysians (Baskaran, 1987). Within this continuum, Manglish represents the more colloquial and highly localized end of Malaysian English, and this fluid switching is not only a practical linguistic tool but also a social and cultural marker. The multilingual nature of Manglish reflects the socio-political and cultural landscape of Malaysia, where linguistic diversity is commonplace. Vollmann and Wooi (2020) described Manglish as a hybrid linguistic form that reflects Malaysia's multicultural identity and dynamic linguistic landscape, and highlights its cultural significance beyond mere linguistic variation. Its

787

prevalence in media outlets such as television shows, advertisements, and social media further underscores its cultural significance and widespread use among various communities. This highlights the relevance of studying Manglish, as it offers critical insights into the linguistic processes of hybridity and the social functions of language in a multilingual society.

Research into Manglish has focused significantly on its lexical innovation, especially its processes of borrowing, calquing (loan translation), and word formation. Scholars have highlighted how borrowing from Malay, Chinese, and Tamil has driven linguistic change in Manglish (Tan, 2009; Baskaran, 2005). Terms like *makan* (eat), *kiasu* (afraid to lose), and *paiseh* (embarrassed) are examples of lexical borrowing that are integrated into English discourse, serving as markers of cultural identity and connecting English with the local linguistic landscape.

In addition to borrowing, calquing is also a prominent feature of Manglish. Calquing involves the direct translation of phrases into English, retaining the local meaning while aligning with English syntax. For example, the Malay phrase *buka lampu* (turn on the light) becomes *open the light* in Manglish, reflecting local meaning in an English structure (Hashim & Tan, 2012). This phenomenon, known as loan translation, illustrates how Manglish speakers creatively adapt English to fit their multilingual environment.

Word formation processes such as compounding, blending, and derivation further contribute to the lexicon of Manglish. Terms like *rojak culture* (mix of cultures) highlight Malaysia's multicultural society, while expressions such as *sohai-fied* (made foolish) illustrate the blending of English and Malay linguistic forms (Baskaran, 2005). These processes demonstrate the morphological productivity of Manglish, generating new words that reflect local meanings and social nuances.

A critical component of Manglish's lexical innovation is the use of discourse markers like *lah*, *lor*, and *meh*, which serve important pragmatic functions (Kuang, 2017). These markers signal social meanings such as solidarity, emphasis, politeness, and informality (Tay et al., 2016). For instance, *lah* is often used to express familiarity and emphasis, as in the sentence *Don't be so serious, lah* (Hashim & Tan, 2012). These markers are integral to the pragmatic dimension of Manglish, marking the speaker's relationship with the listener and reflecting the cultural values of Malaysia's diverse communities.

2.2 Chronological Review of Lexical Innovation in Manglish
2.2.1 Early Works: Borrowing and Nativization (1980s–2000s)

Early studies, such as those by Kachru (1985) and Schneider (2007), laid the foundation for understanding how English adapts to postcolonial contexts. Kachru's (1985) model of World Englishes emphasized the flexibility and nativization of English in non-native settings, where local languages influence the structure and vocabulary of English varieties. Similarly, Schneider's (2007) Dynamic Model of postcolonial Englishes underscored the processes of lexical adaptation and grammaticalization that occur as English becomes integrated into multilingual societies. These frameworks provided the theoretical underpinning for later studies of Manglish, positioning it as a legitimate variety shaped by ongoing language contact.

2.2.2 The Role of Hybridization and Calquing (2000s–2010s)

In the 2000s, research on hybridization in World Englishes gained traction, with scholars such as Buschfeld and Kautzsch (2017) exploring how local languages integrate with English to create hybrid forms. Hybridization in Manglish can be seen in expressions like *lepakable* (suitable for hanging out), where a local element combines naturally with an English suffix to form a new lexical unit with local cultural meaning. This phenomenon reflects the structural integration of local languages with English (Buschfeld & Kautzsch, 2017), suggesting that Manglish is not merely a set of borrowed words but a fully integrated variety that embodies both English and local linguistic forms.

2.2.3 Contemporary Approaches: Word-Formation and Borrowing (2010s–Present)

Recent studies have shifted toward a more comprehensive analysis of internal word-formation processes and their interaction with contact-induced mechanisms. Kalukar et al. (2023) developed a typology of word-formation processes that includes derivation, compounding, and blending, which are all prominent in Manglish. The term *roti-boy* (bread boy), for instance, is a compound blend that combines an English word with a local cultural reference. Similarly, *sohai-fied* (made foolish) demonstrates how derivation and blending work together in Manglish to produce new forms that reflect local experiences and social norms.

While previous research on Manglish has provided valuable insights into its socio-cultural role and lexical innovation, the structural features of these innovations remain underexplored. In particular, the ways in which specific types of lexical innovations are distributed and the socio-cultural factors that influence their creation require further investigation. To fill this gap, this study seeks to address the following research questions and objectives:

Research Questions

RQ1: What types of lexical innovations occur in Malaysian English (Manglish)?

RQ2: How frequently and across what distribution do these innovations occur within the structural categories of word-formation, borrowing and loan translation, and hybridization and localization?

RQ3: What cultural and sociolinguistic factors motivate these structural innovations, and how do they reflect Malaysian identity and localized linguistic creativity?

Research Objectives

RO1: To investigate the types of lexical innovations that occur in Malaysian English (Manglish).

RO2: To analyze the frequency and distribution of these lexical innovations across the structural categories of word-formation, borrowing and loan translation, and hybridization and localization.

RO3: To examine the cultural and sociolinguistic factors that motivate these structural innovations and explain how they reflect Malaysian identity and localized linguistic creativity.

3. Methods

3.1 Corpus Construction

The corpus for the present study was constructed from the book *Manglish: Malaysian English at Its Wackiest!* Authored by Lee Su Kim and Stephen J. Hall. The book contains around 20,000 words across 150 pages and provides authentic examples of Malaysian English expressions (Lee & Hall, 2019). It was selected because it is a widely recognized and accessible resource that documents the distinctive features of Manglish in a systematic and illustrative manner. The book is published in English and reflects current usage practices of Malaysian English in both humorous and descriptive contexts. The text was extracted in full and subsequently cleaned by removing irrelevant content such as headers, footers, and page numbers to ensure consistency. Only the main textual material was retained for analysis. The cleaned text was then converted into a Word file for manual text analysis, with all non-linguistic artifacts excluded. This process ensured that the resulting corpus was both reliable and representative for the purpose of examining Manglish.

3.2 Analytical Framework

This study investigates the structural patterns of lexical innovation in Manglish by building on Kalukar et al. (2023), whose model provides one of the most recent and comprehensive classifications of English word-formation processes, including derivation, inflection, conversion, compounding, blending, abbreviation, acronym, clipping, reduplication, borrowing, and double processes. While this framework effectively captures how new English words are created within the internal morphological system, it remains English-monolingual-centric and thus insufficient to explain the contact-induced innovations that characterize Manglish—a localized variety shaped by the interaction of Malay, Chinese, Tamil, and English. Because Manglish emerges from sustained multilingual contact, its lexical creativity involves not only the internal word-formation processes described by Kalukar et al. (2023) but also contact-driven borrowing, code-mixing, and hybridization that reshape English forms within a multilingual context. To address this limitation, the present study adapts and extends the morphological framework of Kalukar et al. (2023) by incorporating Tan (2009), who outlines loanwords, compound blends, and loan-translation patterns typical of Malaysian English, and Buschfeld & Kautzsch (2017), whose Extra- and Intra-territorial Forces (EIF) model accounts for the hybridization, structural evolution, and localization of English varieties in postcolonial and non-postcolonial contexts. The resulting integrated framework (See Table 1) thus enables a comprehensive analysis of both internal morphological processes and external contact-driven innovations, capturing the full range of structural mechanisms that together define the lexical distinctiveness of Manglish.

Table 1: Structural Taxonomy for Lexical Innovation in Manglish (Adapted from Kalukar et al., 2023; Tan, 2009; Buschfeld & Kautzsch, 2017)

Main Categories	Subtype / Processes	Definitions / Structural Features	Examples (from Manglish)
1. Word-Formation (Internal Morphological Processes)	**Derivation**	Addition of affixes to change word class or meaning.	*datukship, lepakable, stylo*
	Inflection	Affixation marking tense, number, aspect, or degree without changing word class (often used humorously).	*lahs, thankyous*
	Conversion (zero-derivation)	Functional shift of word class without affixation.	*to Google, to message*
	Compounding	Combination of two or more free morphemes into one lexical unit.	*lah-speaker, mosquito biker*
	Blending	Fusion of segments from two words to form a new item.	*bladiful, rainshine, edutainment*
	Abbreviation	Formation using initial letters pronounced individually.	*SMK = Sekolah Menengah Kebangsaan*
	Acronym	Formation using initial letters pronounced as a word.	*MIDA = Malaysian Industrial Development Authority*
	Clipping	Shortening of a longer word without altering meaning or grammatical class.	*sabo, admin, biz, promo*
	Reduplication	Repetition of a word or part of it for emphasis or stylistic effect.	*wait-wait, can-can, relax-lah relax*
	Borrowing (within English)	Adoption of lexical items already circulating *within English* (other English	*dunno, gonna, wanna, kinda, lol, omg*

Main Categories	Subtype / Processes	Definitions / Structural Features	Examples (from Manglish)
		varieties, registers, or online slang), later conventionalized/localized in Manglish use.	
	Double Process	A combination of two or more word-formation mechanisms in one form.	*sohai-fied, frus-case*
2. Borrowing and Loan Translation (Contact-Linguistic Extensions)	**Loanword (direct transfer)**	Direct adoption of lexical items from Malay, Chinese, or Tamil into English discourse.	*curi ayam, kena saman, aiyo, kiasu, kancheong, lepak, makan, tapau*
	Compound blend	Expression combining English and local elements within a single form.	*kopitiam-style, roti-boy, whole jin gang, rojak culture*
	Loan translation (calque)	Literal translation of a local phrase using English morphemes.	*close eye, make dunno, yesterday night, wet market, your head!*
3. Hybridization and Localization (Structural Integration)	/	Combination of English lexical bases with local pragmatic particles or syntactic patterns, reflecting hybrid and localized usage.	*don't like-lah, can meh? so funny lor!*

As shown in Table 1, the analytical framework integrates both internal and external structural processes to account for lexical innovation in Manglish.

The first category, Word-Formation (Internal Morphological Processes), captures the major ways new lexical items are formed or extended within English morphology and usage. Derivation and inflection both involve affixation, but they differ in function: derivation changes meaning or word class (e.g., *datukship, lepakable*), whereas inflection adds grammatical marking without changing class, often in playful or creative forms (e.g., *lahs, thankyous*). Conversion (zero-derivation) shifts word class without affixation, commonly producing verbs from nouns or brand terms (e.g., *to Google, to message*). Compounding and

blending combine lexical bases: compounds join full forms into a single unit (e.g., lah-speaker, mosquito biker), while blends fuse shortened segments from two words (e.g., *bladiful, edutainment*). Abbreviations reduce multi-word expressions to initial letters pronounced individually (e.g., *SMK*), whereas acronyms are pronounced as words (e.g., *MIDA*). Clipping shortens longer words while keeping core meaning and word class (e.g., *sabo, admin, biz, promo*), and reduplication repeats a base wholly or partially for emphasis and interactional effect (e.g., *wait-wait, can-can, relax-lah relax*). In addition, borrowing within English refers to the uptake of lexical items already circulating in other English varieties, registers, or online slang (e.g., dunno, gonna, wanna), which are subsequently conventionalized in Manglish use. Double processes combine multiple mechanisms in a single form (e.g., *sohai-fied*, which integrates a borrowed base with English derivation).

The second category, Borrowing and Loan Translation (Contact-Linguistic Extensions) (following Tan, 2009), focuses on cross-linguistic transfer. Loanwords involve direct transfer of local items or expressions into English discourse (e.g., *curi ayam, kena saman, aiyo, kiasu, kancheong, lepak, makan, tapau*). Compound blends combine English and local elements within one expression, reflecting mixed lexical sourcing (e.g., *kopitiam-style, roti-boy, whole jin gang, rojak culture*). Loan translation (calque) reflects a literal rendering of local pragmatic expressions using English morphemes (e.g., *close eye, make dunno, yesterday night, wet market, your head!*), highlighting indirect transfer through meaning and discourse patterns.

The third category, Hybridization and Localization (Structural Integration) (based on Buschfeld & Kautzsch, 2017), captures innovations where English lexical bases are structurally integrated with local pragmatic particles and localized syntactic patterns, producing distinct Manglish interactional styles (e.g., *don't like-lah, can meh? so funny lor!*).

3.3 Data Analysis

For RQ1, which investigates the types of lexical innovations occurring in Malaysian English (Manglish), the analysis began with AI-assisted lexical screening of the cleaned and tokenized corpus to detect potential new or non-standard lexical items. The AI (ChatGPT 5.2) was guided using a detailed analytical prompt that directed the model to identify, analyze, and categorize each lexical innovation systematically. The exact prompt used for this task was as follows: *"Could you please comprehensively analyze the lexical innovations in the following chapter and accurately categorize them into the following groups: (1) Word-Formation (Derivation, Inflection, Conversion (zero-derivation), Compounding, Blending, Abbreviation, Acronym, Clipping, Reduplication, Borrowing (within English), Double Process); (2) Borrowing and Loan Translation (Loanword (direct transfer), Compound blend, Loan translation (calque)); (3) Hybridization and Localization; and (4) Others? Please also present the categorized items in a table format."* All candidate items identified by the AI were subsequently examined through detailed human verification to ensure linguistic authenticity and to exclude irrelevant or non-lexical tokens. Verified items were then manually validated and classified according to the integrated analytical framework combining Kalukar et al.

(2023) for internal word-formation, Tan (2009) for borrowing and loan translation, and Buschfeld & Kautzsch (2017) for hybridization and localization. This dual-process approach ensured both comprehensive coverage and human analytical reliability.

For RQ2, which explores the frequency and distribution of these innovations, each classified lexical item was counted and recorded according to its structural type. Frequency and proportional distributions were calculated to determine the relative productivity of the three main categories—word-formation, borrowing and loan translation, and hybridization and localization. Descriptive statistics and frequency tables were generated using SPSS (Version 29) to visualize distributional patterns and ensure quantitative accuracy. These numerical results were then interpreted to identify which structural mechanisms were most dominant in Manglish lexical creativity.

For RQ3, which explores the cultural and sociolinguistic factors motivating Manglish lexical innovations, the analysis adopted a qualitative interpretive approach. The classified lexical items identified in RQ1 and RQ2 were examined in relation to their textual, cultural, and pragmatic contexts to uncover how specific forms function as expressions of humor, solidarity, or identity. Each item was further linked to its linguistic or cultural source - Malay, Chinese, or Tamil - to trace possible motivations such as cultural borrowing, discourse pragmatics, or localized identity expression. These findings were then interpreted through the lens of World Englishes scholarship, comparing Manglish patterns with established models of localization and lexical nativization in other postcolonial Englishes. This comparative perspective allowed the study to evaluate how Manglish both aligns with and extends global patterns of contact-induced lexical creativity. Finally, the analysis illuminated how structural innovation in Manglish functions as a marker of Malaysian identity and linguistic belonging, with future research suggested to include speaker interviews for triangulating cultural perceptions and attitudinal meanings.

4. Results and Discussion

4.1 Distribution of Lexical Innovations in Manglish

Figure 1 and Table 1 present the distribution of lexical innovations across various structural categories in Manglish. Based on the results, the most frequent category of lexical innovation is Borrowing and Loan Translation (Contact-Linguistic Extensions), accounting for 45.8% of the total innovations. Within this category, Loanword (direct transfer) (31.5%) > Loan Translation (calque) (12.4%) > Compound Blend (2.0%). The second most frequent category is Hybridization and Localization (Structural Integration), which accounts for 33.1% of the total innovations. The third major category is Word-Formation (Internal Morphological Processes), comprising 21.1% of the innovations. Within this category, Borrowing (within English) (7.8%) > Blending (2.4%) > Reduplication (2.0%) > Derivation (1.6%) & Clipping (1.6%) > Conversion (zero-derivation) (1.3%) & Compounding (1.3%) & Abbreviation (1.3%) & Double Process (1.3%) > Acronym (0.5%).

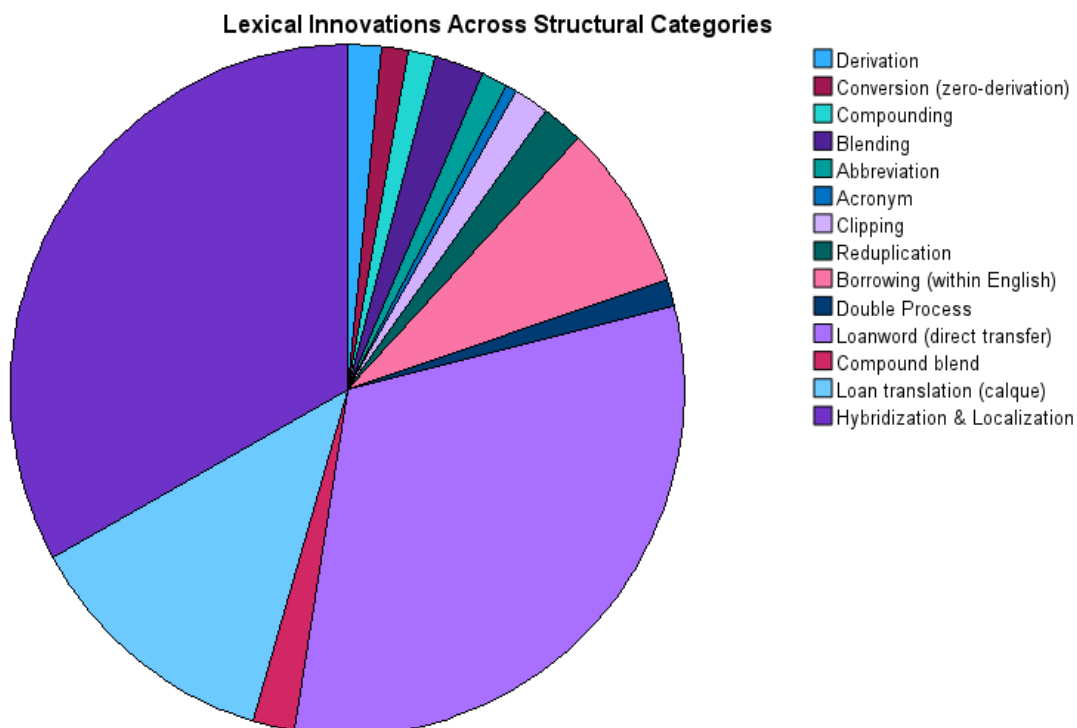**Figure 1: Lexical Innovations Across Structural Categories**

Table 1: Frequency and Percentage of Lexical Innovations Across Structural Categories in Manglish

Main Categories	Subtypes / Processes	Counts	% of Total
1. Word-Formation (Internal Morphological Processes)	**Derivation**	9	1.6%
	Inflection	0	0.0%
	Conversion (zero-derivation)	7	1.3%
	Compounding	7	1.3%
	Blending	13	2.4%
	Abbreviation	7	1.3%
	Acronym	3	0.5%

	Clipping	9	1.6%
	Reduplication	11	2.0%
	Borrowing (within English)	43	7.8%
	Double Process	7	1.3%
	Subtotal (Word-Formation)	116	21.1%
2. Borrowing and Loan Translation (Contact-Linguistic Extensions)	**Loanword (direct transfer)**	173	31.5%
	Compound blend	11	2.0%
	Loan translation (calque)	68	12.4%
	Subtotal (Borrowing & Loan Translation)	252	45.8%
3. Hybridization and Localization (Structural Integration)	/	182	33.1%
	Subtotal (Hybridization & Localization)	182	33.1%
Grand Total	/	550	100%

4.2 Frequency and Sociocultural Drivers of Lexical Innovations in Manglish

The analysis of lexical innovations in Manglish reveals the distribution of various structural categories, with Borrowing and Loan Translation (Contact-Linguistic Extensions) emerging as the most frequent category in the dataset, reflecting long-term contact and everyday mixing among Malay, Chinese varieties, Tamil, and English in Malaysia. This dominance suggests that Manglish grows less through "inventing" entirely new words and more through recycling and recombining multilingual resources that speakers already command. Sustained interaction across communities has thus produced a shared lexical pool that feels locally natural, socially recognizable, and communicatively efficient.

At the heart of this category are direct loanwords, which supply Manglish with compact, high-frequency items for social labeling, evaluation, daily activity, and emotion, as well as

loan translations (calques) and contact-driven semantic transfer into English-shaped expressions. Social labels such as *Ah Beng, Ah Lian, Mat Salleh,* and *towkay* function as quick "social typing" devices. Trait words like *kiasu, kancheong, sombong,* and *geng* evaluate attitudes and personalities with a strong local flavor. Everyday action forms such as *lepak, tapau, belanja, sapu, kena* (be subjected to), and *kantoi* package routine informal practices into short, punchy expressions that fit smoothly into casual speech. Emotional interjections—*Alamak!, Aiyah!, Celaka!*—serve as ready-made reaction markers, signaling surprise, annoyance, shock, or frustration.

These borrowed items come from multiple sources, reflecting Malaysia's multilingual ecology. Many widely used items are Malay in origin (e.g., *habis, leceh, rugi, relek*), especially for everyday evaluation. Chinese varieties such as Cantonese and Hokkien contribute stance- and face-related terms (e.g., *paiseh, lansi, humsup*) that often carry dense pragmatic meanings. Tamil contributes informal address terms used in inter-ethnic peer talk (e.g., *macha*), further enriching Manglish's interpersonal repertoire.

Beyond direct borrowing, Manglish is also productive in loan translation (calquing) and structural semantic extensions, where local idioms, naming habits, and discourse routines are mapped onto English forms while preserving local meanings. Expressions such as *wet market* and *yesterday night* reflect localized naming and time-expression patterns. Others encode interactional strategies and social intent: *close eye* (overlook), *make dunno* (pretend not to know), and *one kind* (a particular type/character) sound English but follow local conceptual patterns. Semantic extensions of English verbs, such as *follow* (ride along) and *send you* (escort you), similarly reflect Malaysian discourse logic.

A further contact-driven mechanism is compound blends, where speakers combine an English structural frame with local elements to create mixed expressions (e.g., *Jinjang Joe, Koochi Rat, Tofu one*), often for vivid characterization, humor, and in-group bonding. These forms keep the accessibility of an English pattern while embedding culturally resonant local cues that immediately signal shared identity—making lexical innovation not random mixing, but socially meaningful stylistic work.

Several factors explain why borrowing, compound blends, and calquing/semantic transfer dominate. Malaysia's intense multilingual contact makes borrowing a natural route for innovation: when languages are used side by side in daily life, speakers often adopt the best-fitting item already available in another language, reflecting the general contact-linguistic tendency to bypass communicative barriers by seeking compromise between their forms of speech (Winford, 2003). Cultural salience further strengthens borrowing because many terms are tightly tied to Malaysian life and lose cultural imagery if replaced (e.g., *mamak stalls*, *nasi lemak*, *ang pow* "red packets"), consistent with Malaysian English lexical features concentrated in culturally embedded domains such as food and festivals (Tan, 2014). These forms offer both efficiency and identity value: they are interactionally economical because meanings are widely shared, and socially meaningful because they signal belonging, local identity, familiarity, and group solidarity. Functionally, this category supports key communicative needs in Manglish. It enables naming and social

typing through labels like *Ah Beng, makcik, samseng,* and *Mat Salleh/towkay*, which sketch recognizable social figures quickly. It provides resources for stance and affect, as in *shiok, suay, teruk,* and (in intensified uses) *habis*, which convey evaluation and emotional coloring. It captures routine actions and informal behaviors through verbs like *lepak* and *cabut*, as well as colloquial figurative verbs such as *goreng, hentam,* and *tembak*, making casual narratives sound locally grounded. It also offers formulaic interjections such as *Alamak!, Aiyah!,* and *Aduh!* for spontaneous reaction. Through tag transfer and semantic extension (and in some cases calquing), it performs pragmatic alignment by mapping local discourse norms into English-shaped expressions—seen in *follow / send you* (ride along / escort)—showing that Manglish innovation is not only lexical, but also interactional and cultural.

Hybridization and Localization (33.1%) shows how Malaysian English is not simply "English with extra words," but English that has been structurally retooled to fit local interactional needs. In this pattern, Malaysian speakers keep an English syntactic backbone while fusing in particles, fixed templates, and TAM-like time/aspect cues in informal interaction. The result is an efficient, highly recognizable style that carries strong identity value: it communicates quickly, manages relationships smoothly, and instantly indexes "Malaysian-ness" in everyday talk.

A prominent mechanism is the use of discourse particles embedded inside English clauses, such as *lah, ah, meh, leh,* and *lor*. These particles function less as "content words" and more as pragmatic tools that adjust tone, stance, and engagement (Tay et al., 2016). For example, "Don't call me Auntie lah!" adds insistence but can also soften the force into a familiar, in-group scolding; "how leh?" turns a problem into a shared concern, inviting response and collaboration. In practice, particles package emphasis, mitigation, warmth, teasing, or skepticism into a single syllable, allowing speakers to manage interpersonal meaning without long explanations. In addition to Manglish, *lah* is also described as one of the most emblematic discourse markers in Colloquial Singaporean English, where it indexes emphasis as well as solidarity, familiarity, and informality (Inharjanto, 2024).

Hybridization also appears through localized constructions and templates that behave like ready-made grammar patterns. Expressions such as *Where got?* operate as a fast existential/negation format ("no such thing / not possible"), while bare affirmative *Can.* works as a complete answer meaning "okay / possible / permitted." Other templates include disbelief frames like *How can?* and counter-challenges like *Why cannot?*, which compress negotiation and argumentation into short, routinized turns. The nominalizer *one*—as in "very blur one" or "can-one ah?"—is another localization that turns descriptions into type-like characterizations ("the kind that is blur") and can even combine with particles to add stance (e.g., skepticism or teasing).

A further hybrid pathway is Malay framing with English content, where Malay grammatical or idiomatic frames carry English "payload". The clearest example is *kena* + English item (e.g., *kena burn, kena firing*), where *kena* functions like a passive/adversative marker that highlights an unwanted or impactful experience, while the English word supplies the event itself. Similarly, Malay chunks can be inserted into English discourse for affective force or

evaluation, such as *tak boleh tahan* embedded within an otherwise English stretch, producing a strongly local rhythm of complaint or emphasis.

Localization is also visible in informal time/aspect marking and reduced forms, especially in casual speech and texting. Forms like *oredi* ("already") and sequences such as "oredi cabut" show how speakers express completion and sequencing in a way that sounds natural in Malaysian interaction, even when it departs from standard written English spelling or morphology.

Another cluster involves procedural and spatial schemas, where English is shaped to deliver compact step-by-step routines for everyday tasks—especially giving directions. Patterns like "roundabout take 3 o'clock / 12 o'clock," "U-turn back," and sequencing strings like "go, go, go" represent a localized "instructional grammar": short imperatives and directional metaphors are chained together to create quick, usable navigation talk. This reflects a broader discourse style in which speakers prioritize practical clarity and rapid sequencing, often using repeated verbs and simplified structures for real-time guidance.

At last, hybridization shows up through code-mixed noun phrases and collocations that blend English with local lexicon to produce vivid evaluation and imagery. Phrases like *paper lama man* combine an English noun with a Malay descriptor and an English discourse add-on; *damn shiok* pairs an English intensifier with a local affect word; and similes like *blur like sotong* use a locally salient image to communicate a very Malaysian kind of humor and social typing. These are not merely isolated borrowings, but localized collocational frames that function as recognizable interactional styles.

Across these examples, the communicative logic is consistent. Pragmatic stance marking is achieved economically through particles that carry emphasis, mitigation, and solidarity. Interactional efficiency is supported by fixed question–answer frames such as "can or not?" and "Where got?", which streamline negotiation, permission-seeking, and disagreement. Procedural discourse becomes concise and stepwise for routines like navigation. Most importantly, there is a clear grammar–pragmatics fusion: English word order remains largely intact, but local discourse norms shape how time, stance, topic, and focus are signaled (often via particles, templates, and TAM proxies rather than standard English morphology). Because these forms are widely recognized, using them functions as identity performance, instantly indexing local belonging and in-group membership while keeping interaction smooth and efficient.

Word-formation in Manglish (21.1%) highlights a different side of lexical innovation: instead of importing material from other languages or reshaping English through contact templates, speakers creatively exploit English's internal resources—borrowing (within English), blending, reduplication, clipping, conversion, derivation, compounding, and localized semantic shift—to build expressions that are vivid, economical, and stylistically "local". Its proportion is lower than contact-based categories mainly because cross-linguistic borrowing is socially and cognitively easier. Loanwords and calques arrive with strong cultural "ready-made" meanings and are immediately recognizable as Malaysian, so they are highly visible identity markers. By contrast, English-internal word-formation and intra-English

borrowing often require more inferencing—listeners must infer the intended shift, blend, or extension—and some formations can sound more like general informal English than distinctly Malaysian English unless they become widely shared conventions. Still, word-formation remains productive because it fills stylistic and pragmatic niches: it is especially useful for humor, rhythmic emphasis, quick stance-taking, teasing, and conversational efficiency.

Before turning to local semantic shifts, it is important to note that this category also includes borrowing (within English), where speakers adopt informal or globally circulating English forms from other English varieties, registers, or online slang and then conventionalize/localize them in Manglish use. A major process here is localized semantic shift (semantic extension of English forms), where ordinary English words are reused with locally conventional meanings and pragmatic force. Items such as *Best!* ("excellent"), *Finish!* ("I'm doomed / I'm done for"), *Gone!* ("ruined"), *Solid!* ("awesome"), and *Terror!* ("amazing") show how evaluation can be intensified through short, punchy English forms that function like interjections or stance markers. Similarly, verbs and nouns are reassigned local meanings that match everyday routines: *fire* (v.) meaning "scold," *chop* (n.) meaning "official stamp," and *fetch* (v.) meaning "pick up someone." These shifts keep familiar English forms but load them with locally shared usage, making them efficient tools for emotion, judgment, and quick storytelling.

Manglish also shows creative blending, where parts of words are merged to compress meaning and produce humorous, memorable forms. Examples such as *bladiful* (*bloody + beautiful*), *rainshine* (rain + shine), and *edutainment* (education + entertainment) illustrate how blends can deliver evaluation or commentary in a single compact unit. Blends often sound playful and expressive, which is why they appear frequently in joking, teasing, and informal narration: they create a "local voice" by turning everyday English materials into stylized, catchy coinages.

Reduplication is another highly recognizable Manglish strategy, using repetition for emphasis, rhythm, or sequencing. Alongside the corpus-attested forms in Table 1 (e.g., *wait-wait, can-can*), expressions like *bluff-bluff* add playfulness and intensification, while time and manner sequences like *next-next Saturday* and *straight-straight* build clarity through rhythmic repetition. Reduplication works well in fast conversation because it signals "extra strength" (more emphatic, more definite, more immediate) without needing longer wording, and it naturally fits the informal, performative tone of Manglish.

Through derivation, speakers attach English or locally adapted affixes to form new adjectives and nouns that match local communicative needs. Words like *heaty* (heat + -y) and *stylo* (style + -o) produce quick evaluative descriptions ("too hot," "stylish/cool"), while *datukship* (datuk + -ship) shows how English morphological patterns can be extended onto local titles to create socially meaningful nouns. Derivation, therefore, strengthens Manglish's evaluative vocabulary and helps speakers label people, styles, and social statuses efficiently.

A related productivity pattern is conversion (zero-derivation), where a word changes grammatical category without added morphology, which suits rapid speech. Using on/off as verbs (e.g., "on the AC") or turning the *horn* into a verb ("horn the bugger") expands the

functional range of existing words while keeping sentences short and action-focused. This is particularly useful in instructions, complaints, and everyday conversation because it streamlines expression and matches the quick tempo of informal talk.

Compounding produces compact, colloquial phrases by combining English elements into (often) locally familiar units. Forms like *somemore* (fused "some more," functioning like "and then / plus also") show lexicalization: the compound becomes a fixed discourse item rather than a transparent phrase. Other compounds, such as *mosquito biker* (and similar descriptive labels in the dataset), likewise pack meaning into a single label, often with a humorous or descriptive edge. Compounding in Manglish keeps English patterns but assigns them local semantic focus, making them useful for quick naming and characterization.

Finally, clipping, abbreviation, and acronyms reflect Manglish's preference for speed and informality, especially in online discourse. Short forms like *sabo* ("sabotage") and the clippings observed in the dataset (e.g., *admin, biz, promo*) increase economy and align with conversational convenience. More generally, abbreviated forms are especially common in digital interaction, where brevity and quick turn-taking are valued. Manglish also shows double processes, where multiple operations combine in one item, as in *frus case* (clipping *frustrated → frus* + compounding with *case*). Such forms demonstrate that, even if word-formation is less dominant than contact-based innovation overall, it remains an active engine for stylistic creativity and pragmatic efficiency in everyday Malaysian English.

5. Conclusion

This study integrated morphological, contact-linguistic, and hybridization frameworks to examine lexical innovation in Manglish using a 20,000-word corpus. The analysis showed a clear distribution across three structural mechanisms: Borrowing & Loan Translation (45.8%), Hybridization & Localization (33.1%), and Word-Formation (21.1%). Together, these patterns demonstrate how English in Malaysia is reshaped into a locally meaningful repertoire that supports naming, stance-taking, interactional management, and cultural indexing. The study's main contribution is a unified taxonomy that bridges internal word-formation with contact-driven processes and constructional hybridity. Quantifying their relative weights shows how contact-induced resources anchor identity quickly, while internal morphology fine-tunes expressivity and economy. The AI-assisted screening by ChatGPT 5.2 plus human coding verification also provides a replicable workflow for small-to-medium corpora of World Englishes.

While the present study offers an empirically grounded snapshot of Manglish innovation, two factors naturally shape how the results should be interpreted and also point to clear next steps. First, because the dataset is a focused 20,000-word, single-source corpus, the frequency profile reported here is best read as a register-sensitive baseline rather than a fully generalizable distribution across all Manglish contexts—especially when compared with spontaneous, multi-party spoken interaction. Second, the integrated taxonomy successfully captures the three dominant mechanisms (borrowing/loan translation, hybridization/localization, and internal word-formation), yet the analysis also surfaced a

small set of boundary cases that are especially typical of rapid informal communication, including colloquial respellings/reductions (e.g., *dunno, dowan, wattaim?*), expressive orthographic stylization (e.g., *Raaaabbishh!*), and nonce playful forms (e.g., *whosed?*). Rather than weakening the framework, these items highlight where Manglish creativity is most active—texting, fast peer talk, and interactional play—and therefore motivate future refinement.

Building on this foundation, subsequent work can expand to larger, multi-genre corpora (e.g., social media, chat logs, podcasts) to stabilize frequency estimates and reduce register bias, and can extend the taxonomy with two targeted micro-categories—Colloquial contraction/respelling and Expressive stylization/elongation—together with clearer operational criteria for distinguishing nonce from conventionalized formations.

References

Baskaran, L. (1987). *Aspects of Malaysian English syntax* (Doctoral dissertation, University of London).

Baskaran, L. M. (2005). *A Malaysian English primer: Aspects of Malaysian English features*. University of Malaya Press.

Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge University Press. https://doi.org/10.1017/CBO9780511845302

Buschfeld, S., & Kautzsch, A. (2017). Towards an integrated approach to postcolonial and non-postcolonial Englishes. *World Englishes*, *36*(1), 104–126. https://doi.org/10.1111/weng.12203

Govindan, I. V., & Pillai, S. (2009). English question forms used by young Malaysian Indians. *The English Teacher*, *38*(1), 74-94.

Hashim, A., & Tan, R. S. K. (2012). Malaysian English. In E. L. Low & A. Hashim (Eds.), *English in Southeast Asia: Features, policy and language in use* (pp. 55–74). John Benjamins.

Inharjanto, A. (2024). The emblematic discourse markers *lah* in Colloquial Singaporean English. *The Journal of English Teaching for Young and Adult Learners, 3*(2), 65–74. https://doi.org/10.21137/jeeyal.2024.3.2.2

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the Outer Circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.

Kalukar, V. J., Erliza, N., & Yahya, M. (2023). A Morphological Analysis of Word Formation Processes in English Posters on Instagram. *Journal of Language Teaching and Research*, *14*(6), 1551-1558. https://doi.org/10.17507/jltr.1406.13

Kirkpatrick, A. (2007). *World Englishes hardback with audio CD: Implications for international communication and English language teaching*. Cambridge University Press.

Kuang, C. H. (2017). The implications of lah, ah, and hah as used by some speakers in Malaysia. *Journal of Modern Languages*, *14*(1), 133-153. Retrieved from https://jml.um.edu.my/index.php/JML/article/view/3801

Lee, S. K., & Hall, S. J. (2019). *Manglish: Malaysian English at its wackiest!* (2nd ed.). Marshall Cavendish Editions. (Orig. 1998).

Pakir, A. (2009). *English in Singapore and Malaysia: Status, features and functions*. Pearson Longman.

Schneider, E. W. (2003). The dynamics of new Englishes: From identity construction to dialect birth. *Language*, *79*(2), 233-281. DOI:10.1353/lan.2003.0136

Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge University Press.

Tan, S. I. (2009). Lexical borrowing in Malaysian English: Influences of Malay. *Lexis. Journal in English Lexicology*, (3). https://doi.org/10.4000/lexis.629

Tan, S. I. (2014). Exploring the Malaysian English Newspaper Corpus for lexicographic evidence. *Kajian Malaysia*, *32*(1), 167-185. Retrived from http://web.usm.my/km/32(Supp.1)2014/KM%2032%20Supp%201%202014%20-%20Art%208(167-185).pdf

Tay, L. C., Chan, M. Y., Yap, N. T., & Wong, B. E. (2016). Discourse particles in Malaysian English: What do they mean? *Bijdragen tot de Taal-, Land- en Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia, 172*(4), 479–509. DOI:10.1163/22134379-17204002

Vollmann, R., & Wooi, S. T. (2020). The sociolinguistic registers of 'Malaysian English'. *Linguistics and Sociology*, 7(1). DOI: 10.47298/cala2020.7-1

Winford, D. (2003). *An introduction to contact linguistics*. Blackwell Publishing.