

Carrying Sarawak Home: Malaysian Malay Code-Mixing in the Indonesian Speech of Returned Migrant Workers from Sambas Regency, West Kalimantan

Riska Wahyuni^{1*}, Lilis Listiyawati², Setya Aristu Pranoto³

^{1,2,3}Politeknik Negeri Pontianak, West Kalimantan, Indonesia

ARTICLE INFO

Article history:

Received: 16 September 2026

Revised: 03 October 2026

Accepted: 18 October 2026

DOI:

10.61100/tacit.v4i2.387

Keywords:

Returned Migrant Workers, Malaysian Malay, Code-Mixing, Sociolinguistics, Social Identity



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

ABSTRACT

Research on migrant language contact often emphasizes interference rather than the relationship between linguistic forms and social meanings after migration. This study examines Malaysian Malay code-mixing among returned Indonesian migrant workers in Sambas Regency. Using a qualitative sociolinguistic approach, data were collected through observation and interviews across three locations. The corpus comprised 25 utterances, including seven observed and eighteen interview-derived examples. Findings revealed lexical insertions, phrase-level expressions, and multiple Malaysian Malay elements within Indonesian speech. Although insertion was prominent, several utterances suggested structural overlap between closely related varieties. Participants reported three motivations: solidarity with fellow returnees, identity-related self-presentation, and residual linguistic habits. English-derived vocabulary and words with divergent meanings also appeared. The study contributes an integrated perspective connecting linguistic structures, migration experiences, and social interpretations. However, the limited corpus and predominance of interview-derived examples restrict conclusions about spontaneous language use and broader social meanings.

ABSTRAK

Penelitian kontak bahasa pada pekerja migran umumnya menekankan interferensi, sementara hubungan antara bentuk linguistik dan makna sosial setelah migrasi belum banyak dikaji secara terpadu. Penelitian ini mengkaji campur kode bahasa Melayu Malaysia dalam tuturan bahasa Indonesia pekerja migran yang kembali ke Kabupaten Sambas. Penelitian menggunakan pendekatan kualitatif sosiolinguistik melalui observasi dan wawancara di tiga lokasi. Data mencakup 25 tuturan, terdiri atas tujuh hasil observasi dan delapan belas hasil wawancara. Temuan menunjukkan penyisipan kata, ungkapan berbentuk frasa, serta penggunaan beberapa unsur Melayu Malaysia dalam satu tuturan. Meskipun penyisipan menonjol, sejumlah tuturan memperlihatkan kemungkinan tumpang tindih struktur kedua bahasa. Informan melaporkan tiga motivasi, yaitu solidaritas, penampilan identitas sosial, dan kebiasaan berbahasa yang bertahan setelah migrasi. Kosakata serapan Inggris dan perbedaan makna leksikal juga ditemukan. Penelitian ini menghubungkan bentuk linguistik, pengalaman migrasi, dan interpretasi sosial. Namun, keterbatasan data, terutama dominasi tuturan wawancara, membatasi kesimpulan mengenai penggunaan bahasa spontan dan makna sosial secara lebih luas dalam masyarakat setempat.

1. INTRODUCTION

Labour migration moves people across borders, but it also moves words. Workers who spend years in another country return with more than savings and stories; they bring back ways of speaking learned on construction sites, plantations, in employers' kitchens, and in shops. Once home, these linguistic habits do not simply disappear. They may surface in greetings, conversations among former co-workers, and ordinary interactions with family members. For communities with a long history of labour migration, an important question is not simply whether returned workers mix languages, but how the linguistic forms acquired abroad become part of everyday communication and what social meanings they carry after returning home.

Sambas Regency provides a relevant setting for examining this phenomenon. Located in northern

* Corresponding author, email address: riskawahyuni83@yahoo.com

West Kalimantan, approximately 227 km from Pontianak, Sambas shares a land border with Sarawak, Malaysia. Formal cross-border movement takes place through the Aruk Cross-Border Post (PLBN Aruk) in Sebunga Village, Sajingan Besar District, while Temajuk Village in Paloh District also represents an important area of local border interaction. Geographical proximity and employment opportunities in Malaysia have encouraged residents of Sambas to seek work abroad, particularly in construction, plantations, domestic work, and services. Some workers remain in Malaysia for several years before returning to their communities. Under Law No. 18 of 2017, these individuals are officially referred to as Indonesian Migrant Workers (*Pekerja Migran Indonesia*, PMI) (Republic of Indonesia, 2017). Their experiences abroad may continue to influence their everyday language use. Expressions such as *Aku pengen pusing-pusing* ('I want to go out for a ride') and *Mengapa gaduh sekali budak-budak ni?* ('Why are these children so noisy?') illustrate how Malaysian Malay elements can appear within Indonesian speech.

Previous research on language contact among Indonesian migrant workers has primarily examined the linguistic influence of Malaysian Malay on Indonesian. Al-Pansori (2014) and Diniarti et al. (2017), for example, identified phonological, morphological, and syntactic interference among returned migrant workers in East Lombok. Their findings indicate that prolonged exposure to Malaysian Malay, social interaction, and established language habits can influence the speech of workers after returning home. Similarly, Sarmin et al. (2023) identified Malaysian Malay elements in Indonesian communication in the Sebatik border area, including words, reduplications, phrases, and clauses. Other studies have demonstrated that code-mixing is influenced by communicative situations and social relationships. These include research involving refugees in Makassar (Hasbi et al., 2025), tourism-related communication (Rijal, 2021), multilingual communities in Lombok (Haerudin & Diniarti, 2021), and ethnic language use among younger speakers (Samatan et al., 2022). Amin & Suyanto (2017) also demonstrated the importance of communicative domains in migrant families' language choices. Collectively, these studies provide important explanations of language contact, linguistic variation, and the social circumstances surrounding multilingual communication.

Beyond linguistic interference, another body of research has examined the relationship between language choice, migration experience, and social identity. Nugroho (2022) found that Indonesian workers returning from Korea used Korean expressions on social media to communicate their overseas experiences and construct particular social identities. Ladegaard (2018) demonstrated how returned domestic workers employed codeswitching to manage emotional alignment and audience relationships when narrating experiences abroad. Hall & Nilep (2015) similarly explained that codeswitching can function as a resource for identity construction in contexts of mobility and globalisation. In the Indonesia-Malaysia border context, Dilah et al. (2023) examined language attitudes and ethnic identification among communities in Temajuk and Telok Melano, while Zakarya Ali et al. (2026) investigated lexical borrowing and morphosyntactic variation in cross-border Malay-speaking communities. These studies suggest that multilingual practices are not merely consequences of language contact but may also reflect belonging, identity, and social relationships.

Despite these contributions, the relationship between the linguistic forms of code-mixing, migrants' experiences abroad, and their social motivations after returning home remains insufficiently integrated in existing research. Studies of Indonesian migrant workers have provided valuable descriptions of interference and linguistic changes associated with exposure to Malaysian Malay (Al-Pansori, 2014; Diniarti et al., 2017). Meanwhile, studies focusing on identity have demonstrated how language choice can communicate belonging, prestige, and personal experience (Ladegaard, 2018; Nugroho, 2022). However, these perspectives have not been sufficiently connected to explain how particular Malaysian Malay words and phrases are used by returned workers in everyday face-to-face interactions and how their meanings vary according to the speakers' experiences, interlocutors, and communicative purposes. This distinction is particularly important in Sambas, where migration to Sarawak creates opportunities for sustained contact between closely related language varieties. The central research gap, therefore, concerns the connection between linguistic structure and social meaning in post-migration communication, rather than the geographical absence of previous studies alone. Furthermore, the structural similarities between Indonesian and Malaysian Malay make lexical insertion relatively accessible, while differences in meaning between similar-looking words may create opportunities for misunderstanding. Examining these characteristics together offers a more comprehensive understanding of how language contact continues to shape communication after migration.

Based on this research gap, the present study examines how Malaysian Malay code-mixing appears in the Indonesian speech of returned migrant workers from Sambas Regency and how its use relates to their migration experiences and social motivations. Specifically, the study addresses two research questions: (1)

What linguistic forms of Malaysian Malay code-mixing occur in the Indonesian speech of returned migrant workers from Sambas Regency? (2) What social and linguistic factors influence the use of these forms after the workers return to their home communities? Through these questions, the study seeks to connect the structural characteristics of code-mixing with the communicative contexts in which they occur.

The novelty of this study lies in integrating the analysis of linguistic forms with the social meanings of migration experience in the everyday speech of returned migrant workers. Rather than treating Malaysian Malay elements primarily as interference with Indonesian, the study examines how word-level and phrase-level insertions may serve different communicative purposes, including solidarity among former migrant workers, the display of overseas work experience as a form of social identity, and the continuation of linguistic habits acquired abroad. This approach extends previous interference-oriented studies by showing how linguistic forms can be interpreted alongside speakers' relationships, intentions, and migration histories. The study also examines English-derived vocabulary encountered through Malaysian Malay and *false friends* between Indonesian and Malaysian Malay, highlighting how contact between closely related varieties may facilitate linguistic integration while creating potential semantic ambiguity. By bringing these dimensions together, the study contributes an integrated sociolinguistic perspective that connects language contact, post-migration identity, and everyday communication in migrant-sending communities.

2. THEORETICAL FRAMEWORK AND HYPOTHESES

2.1. Bilingualism and Language Contact

Language contact occurs when speakers of different languages or language varieties interact regularly, creating opportunities for linguistic influence and multilingual practices (Weinreich, 1953). Bilingualism is closely associated with this process, although the ability to use more than one language does not necessarily develop equally across all communicative situations. Bhatia & Ritchie (2013) explain that bilingual language practices are shaped by social environments, communicative needs, and relationships between speakers. Individuals may acquire additional linguistic resources through family interaction, education, employment, and migration. These experiences influence not only the languages speakers know but also how they select and combine linguistic elements in everyday communication.

In Indonesia, many speakers use regional languages alongside Indonesian, which serves as the national language and an important medium of communication across ethnic communities. For residents of Sambas Regency, employment in Sarawak creates further opportunities for contact with Malaysian Malay, including regional varieties used in workplaces and local communities. Returned migrant workers may therefore possess linguistic repertoires that include Indonesian, Sambas Malay or Dayak languages, and elements acquired through communication in Malaysia. This multilingual background provides the context for examining why certain Malaysian Malay expressions remain present in their speech after returning home.

2.2. Code-Mixing: Forms and Types

Code-mixing refers to the use of linguistic elements associated with one language or variety within an utterance that also contains elements of another. In Indonesian sociolinguistic studies, code-mixing has commonly been classified according to the linguistic units involved, including words, phrases, clauses, idiomatic expressions, and hybrid forms (*baster*) (Suwito, as cited in Herawati, 2016). This classification is useful for identifying the grammatical forms of mixed elements and describing their occurrence within everyday utterances.

Muysken (2000) offers a complementary explanation by distinguishing three processes of code-mixing: insertion, alternation, and congruent lexicalization. Insertion involves incorporating lexical items or constituents associated with one language into a structure associated with another. Alternation involves switching between grammatical structures associated with different languages, while congruent lexicalization occurs when linguistic varieties share grammatical structures that allow lexical elements from either variety to appear within the same construction. These distinctions are particularly relevant to Indonesian and Malaysian Malay because their structural similarities can make the boundaries between varieties difficult to identify.

In this study, Muysken's (2000) typology provides the primary framework for examining the structural characteristics of code-mixing. The analysis considers whether Malaysian Malay elements occur as words, phrases, or other linguistic units and whether their distribution is consistent with insertion or suggests alternative mixing processes. Importantly, identifying a particular linguistic form does not, by itself,

establish the speaker's motivation for using it. Structural classification explains how language elements appear in an utterance, while understanding why they are used requires additional attention to social context and speakers' accounts.

2.3. Why Speakers Mix: Social Motivations

Code-mixing is influenced by several factors involving speakers, their relationships, and the situations in which communication occurs. Bhatia & Ritchie (2013) identify participant relationships, situational conditions, message-related considerations, and language attitudes as important dimensions of bilingual language practices. Kachru's distinction, as summarised by Herawati (2016), also considers internal and external aspects of language mixing, including the presence of linguistic resources within a community and speakers' linguistic competence and psychological considerations. These perspectives suggest that code-mixing cannot be understood solely through grammatical classification.

Gumperz (1982) provides an interactional perspective through the distinction between *we-code* and *they-code*. The concept of *we-code* refers to linguistic choices associated with informal relationships, shared experiences, and group belonging, while *they-code* is associated with more distant or institutionally defined relationships. This distinction helps explain how speakers may use particular linguistic resources to express familiarity and solidarity. Among returned migrant workers, Malaysian Malay expressions may function as signals of shared experience when speakers interact with others who have also worked in Malaysia.

Myers-Scotton (1993) approaches language choice through the markedness perspective, which considers how speakers use linguistic choices to signal or negotiate social relationships. An unmarked choice is consistent with the expectations of a particular interaction, while a marked choice may communicate a different relationship or social position. In the context of returned migrant workers, Malaysian Malay expressions may be familiar and socially expected among former co-workers. In other situations, their use may become more noticeable and may be interpreted as a way of displaying overseas work experience.

Gumperz's (1982) and Myers-Scotton's (1993) perspectives complement one another but address different aspects of social meaning. Gumperz helps explain how language choice can communicate group belonging, whereas Myers-Scotton provides a framework for interpreting whether that choice confirms or negotiates expected social relationships. Neither perspective assumes that a particular word or phrase necessarily expresses solidarity or social status. Such interpretations must be supported by the interactional context and, where available, the speakers' own explanations.

2.4. Code-Mixing, Identity, and Migration

Migration can reshape speakers' linguistic repertoires through prolonged interaction with communities that use different languages or varieties. Linguistic resources acquired abroad may remain part of everyday communication after migrants return home. However, the continued use of these resources does not necessarily reflect a single motivation. Depending on the interaction, speakers may use familiar expressions to maintain social relationships, communicate aspects of their experiences, or simply draw on established linguistic habits.

Hall & Nilep (2015) explain that codeswitching can serve as a resource for constructing and negotiating identity under conditions of mobility and globalisation. Nugroho (2022) found that returned Indonesian migrant workers used Korean expressions on social media in ways associated with their overseas experiences and social identity. Similarly, Ladegaard (2018) demonstrated how codeswitching in returned domestic workers' narratives could contribute to emotional alignment and audience positioning. These studies indicate that language practices associated with migration can carry meanings beyond their immediate linguistic content.

For returned migrant workers in Sambas, Malaysian Malay code-mixing may therefore be examined from two related perspectives. The first concerns the linguistic forms retained through contact with Malaysian Malay during employment abroad. The second concerns how these forms are used and interpreted within social relationships after returning home. Connecting these perspectives makes it possible to examine code-mixing as both a linguistic practice shaped by migration and a potential resource for expressing post-migration identity. Nevertheless, social meanings cannot be inferred from linguistic forms alone; they must be interpreted in relation to the available empirical evidence.

2.5. Indonesian and Malaysian Malay: Shared Roots, Divergent Lexicons

Indonesian and Malaysian Malay belong to the Austronesian language family and share historical roots in Malay. Although they have developed through different national, educational, and sociolinguistic contexts, the two varieties retain substantial similarities in vocabulary and grammatical structure. Indonesian has been influenced by several languages, including Sanskrit, Arabic, Dutch, and English, while Malaysian Malay has also developed through contact with English and regional Malay varieties (Al-Pansori, 2014).

These similarities are important for understanding code-mixing among returned migrant workers. Shared grammatical patterns may allow Malaysian Malay lexical elements to appear within Indonesian utterances without substantial changes to sentence structure. At the same time, linguistic similarity does not guarantee identical meanings. For example, *percuma* commonly means 'free of charge' in Malaysian Malay but 'in vain' in Indonesian. Such expressions are commonly described as *false friends*, because similar forms may carry different meanings across language varieties.

The relationship between these varieties presents two analytical considerations. First, their structural similarities may make it difficult to distinguish insertion from congruent lexicalization in some utterances. Second, differences in lexical meaning may create potential ambiguity, particularly when speakers and listeners do not share the same linguistic experience. However, the existence of false friends does not automatically demonstrate that misunderstanding has occurred. Actual misunderstanding would require evidence from listeners or observed interaction. This study therefore treats semantic divergence as a potential communicative issue rather than an established consequence of code-mixing.

2.6. Conceptual Framework

The conceptual framework integrates three complementary analytical perspectives to examine Malaysian Malay code-mixing among returned migrant workers. Muysken's (2000) typology addresses the structural dimension by identifying how linguistic elements from different varieties occur within utterances. Gumperz's (1982) *we-code* concept addresses the interactional dimension by explaining how language choice may communicate shared experience and group solidarity. Myers-Scotton's (1993) markedness perspective addresses the social positioning dimension by examining how linguistic choices may confirm or negotiate expected relationships between speakers.

These perspectives are connected through the experience of migration. Prolonged contact with Malaysian Malay during employment in Sarawak provides opportunities for workers to acquire and repeatedly use particular linguistic elements. After returning to Sambas, these elements may remain available within their linguistic repertoires. Their subsequent use occurs in different interactional settings, including conversations with fellow returnees, family members, and people without similar migration experiences.

The framework does not assume that particular social motivations necessarily produce particular forms of code-mixing. Instead, linguistic forms and social motivations are examined as related but analytically distinct dimensions. Muysken's typology is used to classify the forms and processes observed in the utterances. Gumperz's and Myers-Scotton's perspectives are then used to interpret possible social meanings by considering the interlocutors, communicative situations, and explanations provided by informants. This approach allows the same linguistic element to be interpreted differently across interactions without assuming a fixed relationship between its grammatical form and social function.

The empirical material examined in this study consists of 25 utterances containing Malaysian Malay elements, collected through observation and interviews. The analysis identifies word-level and phrase-level mixing and considers three reported or interpreted motivations: solidarity, status display, and residual linguistic habit. These categories are examined in relation to the available evidence rather than treated as pre-determined causal mechanisms. In particular, solidarity and status display are interpreted through interactional and identity-related perspectives, while residual habit is considered in relation to continued lexical familiarity after prolonged language contact.

Figure 1 presents the framework by distinguishing the study's empirical focus from its theoretical interpretation. Migration experience and language contact provide the sociolinguistic context. Linguistic forms and interactional conditions constitute separate dimensions of analysis. Muysken's typology guides structural classification, while Gumperz's and Myers-Scotton's perspectives guide the interpretation of social meaning. The framework therefore represents an integrated analytical approach rather than a tested causal sequence.

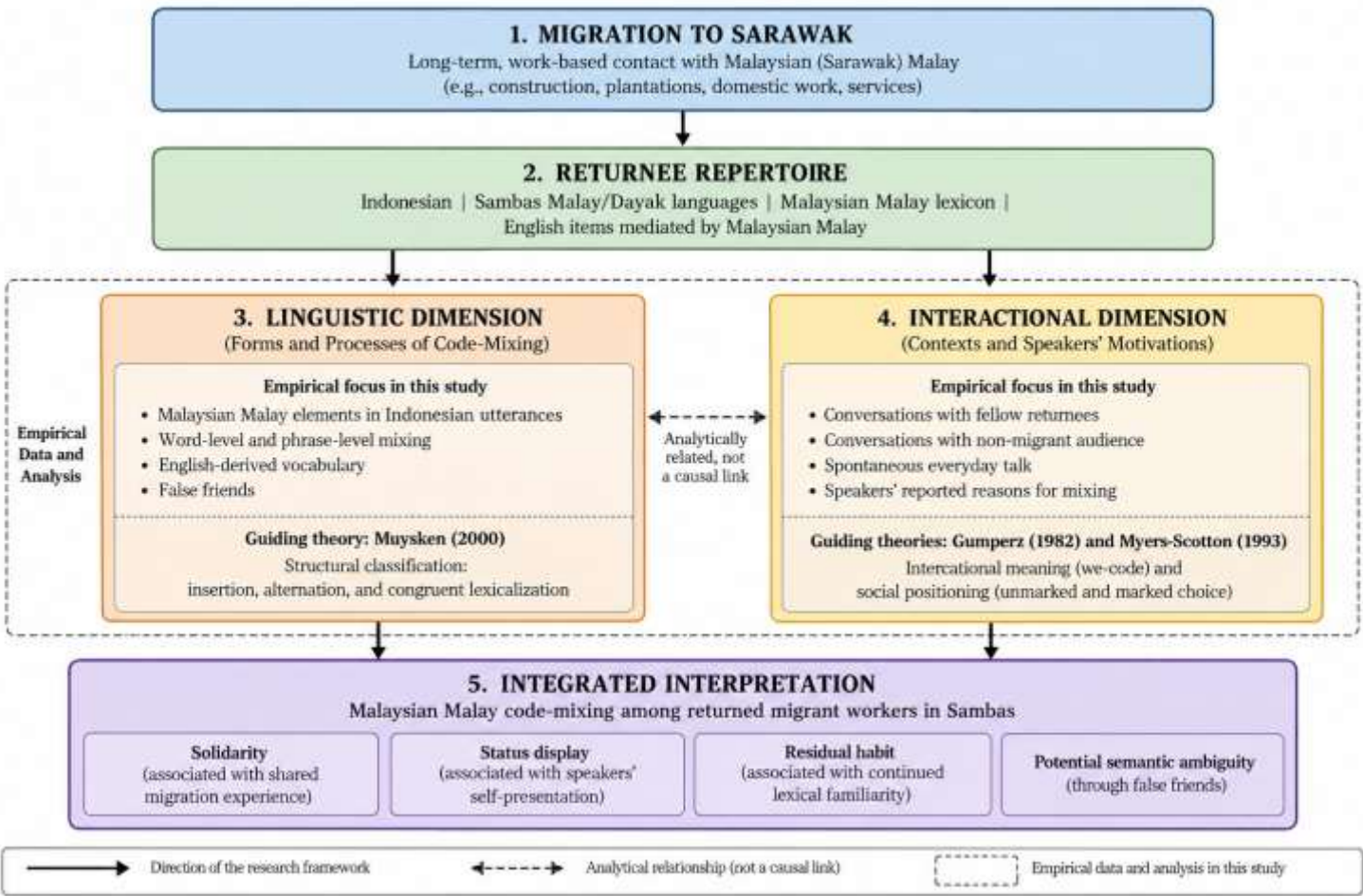


Figure 1. Integrated Analytical Framework of Malaysian Malay Code-Mixing among Returned Migrant Workers

3. RESEARCH METHOD

This study employed a qualitative descriptive design with a sociolinguistic approach to examine Malaysian Malay code-mixing in the Indonesian speech of returned migrant workers from Sambas Regency, West Kalimantan. The approach combined the analysis of linguistic forms with an examination of the social contexts and motivations associated with their use. A qualitative design was appropriate because the study aimed to describe the forms of code-mixing and understand participants' explanations of their language practices rather than measure the prevalence of code-mixing among returned migrant workers. The research was conducted from May to September 2026 in three locations: Sempadian Village, Tekarang District; Mensungai Village, Teluk Keramat District; and Sajingan Besar District.

Informants were selected through purposive sampling based on their previous employment experience in Malaysia and their return to Sambas Regency. The selection focused on individuals who could provide information about Malaysian Malay expressions acquired during migration and their continued use in everyday communication. Informants from Sempadian and Mensungai were identified as Sambas Malay speakers, while those from Sajingan Besar included Dayak speakers. Their migration experiences varied in duration. For example, during an observed household interaction in Sempadian, the husband had worked in Malaysia for five years, his wife for three years, and a visiting guest for three years. These differences in linguistic and migration backgrounds provided relevant contexts for examining language contact after return.

Data were collected through observation, audio recording, note-taking, and semi-structured interviews. Observation employed the listening (*simak*) technique described by Sudaryanto (1993). The researchers documented conversations involving returned migrant workers and identified utterances containing Malaysian Malay elements. Audio recordings and field notes were used to preserve the linguistic expressions

and available information about their communicative contexts. Semi-structured interviews explored the Malaysian Malay vocabulary participants continued to use, situations in which they used particular expressions, their employment experiences in Malaysia, and their explanations for mixing Malaysian Malay with Indonesian.

The study distinguished between naturally occurring utterances and interview-derived examples. Of the 25 code-mixed utterances analysed, seven were obtained through observation and eighteen through interviews. Observational data provided examples of linguistic expressions used in identifiable social interactions. Interview data contributed additional expressions that participants produced or recalled when discussing their language practices. These sources were analysed together to describe the range of code-mixing forms, while their different characteristics were considered when interpreting the findings. In particular, interview-derived expressions were not automatically treated as evidence of spontaneous language use.

Data analysis followed the procedures of data condensation, data display, and conclusion drawing and verification proposed by Miles et al. (2014). During data condensation, the researchers reviewed the documented utterances and selected expressions containing Malaysian Malay elements. The expressions were then transcribed, translated where necessary, and compared with their Indonesian equivalents. During data display, the utterances were organised according to linguistic forms, grammatical categories, and code-mixing processes. The classification considered word-level and phrase-level elements as well as Muysken's (2000) distinction between insertion, alternation, and congruent lexicalization.

The analysis also examined the social meanings associated with code-mixing. Participants' explanations were considered alongside the available interactional contexts to identify recurring themes concerning shared migration experience, social identity, and continued linguistic habits. Gumperz's (1982) *we-code* concept and Myers-Scotton's (1993) markedness perspective guided the interpretation of these themes. Linguistic forms and social motivations were examined as related but distinct dimensions. Accordingly, the presence of a particular Malaysian Malay expression was not considered sufficient evidence of solidarity or status display without supporting contextual information or participant explanations.

To strengthen the credibility of the interpretations, the analysis compared linguistic examples obtained through observation and interviews and examined their consistency across the three research locations. The researchers considered whether the identified linguistic forms corresponded with the meanings reported by participants and whether the available conversational contexts supported the proposed social interpretations. This comparison helped distinguish directly documented language use from participants' accounts of their linguistic habits. The analysis did not assume that all reported motivations could be independently confirmed through observation.

The study also recognised several methodological limitations. First, the corpus consisted of 25 utterances, including eighteen interview-derived examples, and therefore could not establish how frequently particular forms occurred in spontaneous everyday communication. Second, the interpretation of social motivations relied partly on participants' explanations, which may not fully capture the intentions associated with every utterance. Third, the study did not systematically include non-migrant listeners, limiting conclusions about how Malaysian Malay expressions were understood by audiences without migration experience. These limitations were considered when interpreting the findings and formulating the study's conclusions.

4. DATA ANALYSIS AND DISCUSSION

4.1. Overview: Structural Patterns of Malaysian Malay Code-Mixing

The study identified 25 utterances containing Malaysian Malay elements in the Indonesian speech of returned migrant workers from Sambas Regency. The data were obtained through observation and interviews in Sempadian Village, Mensungai Village, and Sajingan Besar District. Of these utterances, seven were documented through observation and eighteen were obtained from interviews. The analysis focused on identifying the linguistic elements associated with Malaysian Malay, examining their grammatical positions, and interpreting their use in relation to the available communicative contexts.

The classification was reassessed by distinguishing the grammatical form of each mixed element from the process through which it appears within an utterance. Following the linguistic-unit classification discussed in Section 2, individual elements were examined as words, phrases, or other grammatical constituents. Muysken's (2000) typology was then applied to distinguish insertion, alternation, and congruent lexicalization. Insertion was identified when a Malaysian Malay lexical item or constituent appeared within a recognisable Indonesian grammatical frame. Alternation was considered when the data suggested a shift

between distinct grammatical structures, while congruent lexicalization was considered when the shared grammatical patterns of Indonesian and Malaysian Malay made the structural boundary between the varieties difficult to determine.

Four criteria guided the classification. First, the analysis identified all potentially relevant Malaysian Malay elements within each utterance rather than selecting only one representative word. Second, the grammatical boundaries of each element were examined to determine whether the utterance contained a single constituent or several separate elements. Third, the surrounding structure was assessed to identify whether Indonesian provided a sufficiently recognisable grammatical frame. Fourth, cases involving shared vocabulary, overlapping grammatical structures, or non-contiguous Malaysian Malay elements were treated cautiously because these features may complicate the distinction between insertion and congruent lexicalization.

The reassessment revealed that many utterances contain identifiable Malaysian Malay lexical items within an Indonesian or locally shared grammatical structure. However, the presence of multiple elements in several utterances prevents the entire corpus from being classified unambiguously as single-word or phrase insertion. Utterance (7), for example, contains both *jom* and *bandar*, while utterances (8) and (12) contain *nak* alongside another Malaysian Malay lexical item. Similarly, utterances (21) and (24) contain more than one Malaysian Malay element distributed across different grammatical positions. These examples require a distinction between the classification of individual constituents and the classification of the complete utterance.

The findings therefore indicate that insertion is a prominent structural pattern in the corpus, while several utterances also display characteristics that may be consistent with congruent lexicalization. This interpretation is consistent with Muysken's (2000) observation that structurally related languages may share grammatical patterns that complicate the identification of clear language boundaries. The data do not provide sufficient evidence to establish alternation as a separate recurring pattern. Nevertheless, the possibility of structural overlap means that an exclusively insertion-based classification would be too restrictive.

Table 1 presents the 25 utterances, their Malaysian Malay elements, and the revised structural interpretation. The table distinguishes the primary linguistic elements from additional elements and identifies cases requiring cautious classification.

Table 1. Malaysian Malay Code-Mixing in the Indonesian Speech of Returned Migrant Workers

No.	Utterance	Malaysian Malay elements	Revised structural interpretation	Source
1	Tolong ambilkan <i>tuala</i> .	<i>tuala</i> (towel)	Single lexical insertion: noun	Observation, Sempadian
2	Bagus benar <i>selipar</i> -mu.	<i>selipar</i> (sandals)	Lexical insertion with possessive attachment	Observation, Sempadian
3	Banyak juga <i>pusak</i> -mu.	<i>pusak</i> (cat)	Single lexical insertion: noun	Observation, Sempadian
4	Dahulu, aku bekerja di Malaysia sering diminta <i>overtime</i> .	<i>overtime</i> (overtime work)	Single lexical insertion: noun	Observation, Sempadian
5	Aku pergi ke <i>kedai</i> ya.	<i>kedai</i> (shop)	Single lexical insertion: noun	Observation, Sempadian
6	<i>Kasut</i> baru ya?	<i>kasut</i> (shoes)	Lexical insertion within a short, structurally shared expression	Observation, Sempadian
7	<i>Jom</i> , kita pergi ke <i>bandar</i> .	<i>jom</i> (come on); <i>bandar</i> (town)	Multiple lexical insertions; interjection and noun	Interview, Mensungai
8	Saya <i>nak</i> pergi beli <i>bag</i> .	<i>nak</i> (want); <i>bag</i> (bag)	Multiple lexical insertions; modal and noun	Interview, Mensungai
9	Saya membeli <i>kerete</i> baru.	<i>kerete</i> (car)	Single lexical insertion: noun	Interview, Mensungai
10	Berapa <i>minit</i> kita sampai ke rumah Erni?	<i>minit</i> (minute)	Single lexical insertion: noun	Interview, Mensungai
11	Bagus <i>sangat</i> <i>basikal</i> -mu.	<i>sangat</i> (very); <i>basikal</i> (bicycle)	Noun insertion; <i>sangat</i> requires contextual verification as a shared lexical item	Interview, Mensungai
12	Aku <i>nak</i> ke <i>tandas</i> .	<i>nak</i> (want); <i>tandas</i> (toilet)	Multiple lexical insertions; modal and noun	Interview, Mensungai

13	<i>Enjin</i> motor aku rusak.	<i>enjin</i> (engine)	Single lexical insertion: noun	Interview, Mensungai
14	Saya <i>gostan</i> -kan mobil Bapak.	<i>gostan</i> (reverse)	Lexical insertion with morphological integration	Interview, Sajingan Besar
15	Kita salah jalan ini. <i>Pusing balik</i> ya!	<i>pusing balik</i> (turn back)	Verb phrase; possible insertion within a shared grammatical structure	Interview, Sajingan Besar
16	Silakan <i>dicuba</i> maka-nannya, Ibu.	<i>cuba</i> (try)	Lexical insertion with morphological integration	Observation, Sempadian
17	Aku mendapatkan barang ini secara <i>percuma</i> .	<i>percuma</i> (free of charge)	Lexical insertion involving semantic divergence	Interview, Mensungai
18	Anak kecil ini <i>comel</i> sekali.	<i>comel</i> (cute)	Single lexical insertion: adjective	Interview, Mensungai
19	Perbaiki rumah ini <i>senang</i> jika punya uang.	<i>senang</i> (easy)	Shared-form lexical item with divergent meaning	Interview, Mensungai
20	Aku pengen <i>pusing-pusing</i> .	<i>pusing-pusing</i> (go around)	Reduplicated verbal expression; possible insertion	Interview, Mensungai/Sempadian
21	Saya <i>nak</i> <i>berkahwin</i> selepas pulang ke Sambas.	<i>nak berkahwin</i> (want to marry); <i>selepas</i> (after)	Multiple non-contiguous elements; possible structural overlap	Interview, Mensungai/Sempadian
22	Aku lupa membawa <i>bag beli-belah</i> dari rumah.	<i>bag beli-belah</i> (shopping bag)	Noun phrase insertion	Interview, Mensungai
23	Bapak ingin pergi ke <i>panggung wayang</i> gambar di Singkawang.	<i>panggung wayang</i> gambar (cinema)	Noun phrase insertion	Interview, Mensungai
24	Mengapa <i>gaduh</i> sekali <i>budak-budak</i> ni?	<i>gaduh</i> (noisy); <i>budak-budak ni</i> (these children)	Multiple non-contiguous elements; possible congruent lexicalization	Interview, Sajingan Besar
25	Perbuatan pemuda itu sungguh <i>tak elok</i> .	<i>tak elok</i> (not good)	Adjective phrase insertion	Interview, Sajingan Besar

4.2. Nouns: The Vocabulary of Daily Life Abroad

Nouns constitute a substantial part of the documented Malaysian Malay elements. They include vocabulary associated with household activities, transportation, workplaces, and everyday social interaction. These lexical items reflect domains in which returned migrant workers were likely to encounter Malaysian Malay during their employment and residence in Sarawak.

Utterances (1)–(6) were documented through observation in Sempadian Village. During Eid al-Adha 2026, a household received visitors from the family of their child's prospective spouse. The host had previously worked in Malaysia for five years, his wife for three years, and a visiting guest for three years. During the interaction, the host asked his wife:

(1) *Tolong ambulkan tuala.*

'Please bring me the towel.'

The visiting guest also commented on the household's cats:

(3) *Banyak juga pusak-mu.*

'You have quite a lot of cats.'

Another expression concerned the visitor's sandals:

(2) *Bagus benar selipar-mu.*

'Your sandals are really nice.'

These utterances demonstrate the presence of Malaysian Malay and Sarawak Malay lexical elements within ordinary household communication. The expressions occurred in an interaction involving individuals who had previously worked in Malaysia. However, the linguistic forms alone do not establish whether the speakers deliberately used these expressions to communicate solidarity or whether they reflected familiar language habits.

Additional observational data involved two women who had become friends while working in Malaysia. Their conversation included references to employment experiences and everyday activities:

(4) *Dahulu, aku bekerja di Malaysia sering diminta overtime.*

'When I worked in Malaysia, I was often asked to work overtime.'

(6) *Kasut baru ya?*

'New shoes, huh?'

Utterances (7)–(13) were obtained through interviews in Mensungai Village. They included nouns referring to places, transportation, household objects, and everyday activities. Examples include *bandar* ('town'), *kerete* ('car'), *basikal* ('bicycle'), *tandas* ('toilet'), and *enjin* ('engine').

The presence of these words suggests that everyday vocabulary remains accessible after migrants return home. This interpretation is consistent with Muysken's (2000) discussion of lexical insertion. Nevertheless, the interview-derived examples demonstrate participants' reported linguistic repertoires rather than the frequency with which these expressions occur in spontaneous conversation.

Several utterances also contain elements beyond the nouns initially identified. In (7), *jom* appears alongside *bandar*, while (8) and (12) contain *nak* together with *bag* and *tandas*, respectively. These examples show that the analysis must consider the complete utterance rather than assigning its structural classification solely on the basis of one selected noun.

4.3. Verbs and Adjectives: Morphological Integration and Semantic Divergence

The verbal and adjectival elements in the corpus provide further evidence of the relationship between Malaysian Malay and Indonesian. Some expressions show morphological integration, while others demonstrate differences in lexical meaning despite similarities in form.

In utterance (14), the lexical element *gostan* appears with the suffix *-kan*:

(14) *Saya gostan-kan mobil Bapak.*

'I'll reverse your car, Sir.'

Similarly, utterance (16) contains *cuba* with the prefix *di-*:

(16) *Silakan dicuba makanannya, Ibu.*

'Please try the food, Ma'am.'

These examples illustrate how lexical elements associated with Malaysian Malay can appear within morphological patterns also found in Indonesian. However, because the two varieties share related grammatical structures, the presence of an affix should not automatically be interpreted as evidence that the entire construction belongs exclusively to Indonesian.

Other expressions demonstrate semantic divergence between the two varieties:

(15) *Kita salah jalan ini. Pusing balik ya!*

'We took the wrong road. Turn back!'

(17) *Aku mendapatkan barang ini secara percuma.*

'I got this item for free.'

(19) *Perbaiki rumah ini senang jika punya uang.*

'Repairing this house is easy if you have money.'

The words *pusing*, *percuma*, and *senang* illustrate how similar lexical forms may carry different meanings depending on the language variety and communicative context. In Indonesian, *pusing* commonly refers to dizziness, *percuma* may mean 'in vain', and *senang* commonly expresses happiness or pleasure. In Malaysian Malay, the meanings intended in these utterances differ.

Similar issues arise with *bandar*, *kerete*, and *budak*. These examples demonstrate the potential for semantic ambiguity when closely related varieties are used within the same communicative setting.

The findings establish the presence of lexical items with divergent meanings, but they do not demonstrate that misunderstanding actually occurred. Since non-migrant listeners were not systematically interviewed, any claim concerning their interpretations must remain provisional. The analysis therefore identifies potential communicative ambiguity rather than confirmed communication failure.

4.4. Phrases and Multiple Insertions: Reconsidering Structural Boundaries

The phrase-level data require particular attention because several utterances contain multiple Malaysian Malay elements that do not form a single continuous grammatical constituent.

Utterance (21) provides an example:

(21) *Saya nak berkahwin selepas pulang ke Sambas.*

'I want to get married after returning to Sambas.'

The expression *nak berkahwin* functions as a verbal construction, while *selepas* introduces a temporal relationship. These elements occur in different grammatical positions. They should therefore not be classified

automatically as one continuous verb phrase insertion. A more cautious interpretation is that the utterance contains multiple Malaysian Malay elements within a structure shared by Indonesian and Malaysian Malay.

A similar issue appears in utterance (24):

(24) *Mengapa gadoh sekali budak-budak ni?*

'Why are these children so noisy?'

Here, *gadoh* functions as a descriptive element, while *budak-budak ni* forms a nominal expression. The two elements are separated by *sekali*. Their distribution does not support treating the entire sequence as a single adjective phrase. Instead, the utterance contains multiple elements whose arrangement may reflect the structural compatibility of the two varieties.

Other examples show more clearly identifiable phrase-level constituents:

(22) *Aku lupa membawa bag beli-belah dari rumah.*

'I forgot to bring the shopping bag from home.'

(23) *Bapak ingin pergi ke panggung wayang gambar di Singkawang.*

'Father wants to go to the cinema in Singkawang.'

(25) *Perbuatan pemuda itu sungguh tak elok.*

'That young man's behaviour is really not good.'

In these cases, the relevant expressions can be analysed as noun or adjective phrases occurring within the surrounding utterance.

The distinction is theoretically important. Muysken (2000) explains that insertion and congruent lexicalization represent different structural processes. Insertion involves identifiable constituents associated with one variety appearing within another grammatical frame. Congruent lexicalization, by contrast, involves the use of lexical elements from different varieties within a shared grammatical structure.

The present corpus suggests that insertion and structural overlap may coexist in the speech of returned migrant workers. However, the limited conversational context, particularly for interview-derived utterances, prevents definitive classification of every ambiguous example. The analysis therefore identifies possible congruent lexicalization without claiming that it has been conclusively demonstrated.

This finding extends the discussion of Malay–Indonesian language contact by showing that grammatical similarity complicates the classification of mixed utterances. Rather than assuming that every Malaysian Malay element represents a straightforward insertion, the analysis recognises the possibility of multiple constituents and shared structural patterns.

4.5. English through Malaysian Malay: Layered Borrowing

Several utterances contain vocabulary associated with English that is also used in Malaysian Malay. These include *overtime*, *bag*, *minit*, *basikal*, *enjin*, *gostan*, and *selipar*. Their presence reflects the multilingual character of the linguistic resources encountered by migrant workers in Malaysia.

For example, utterance (14) contains *gostan*, a lexical form associated with the English expression 'go astern', combined with the suffix *-kan*. Utterance (22) contains *bag beli-belah*, combining an English-derived lexical element with a Malay expression.

These examples illustrate the possibility of layered language contact, in which lexical resources associated with English circulate through Malaysian Malay and appear in the speech of returned migrant workers.

However, the presence of English-derived vocabulary does not independently establish the precise route through which each participant acquired it. The interpretation that these elements were encountered through Malaysian Malay is consistent with the migration context, but the available data do not provide a complete linguistic history of individual words.

This observation remains relevant to the study of cross-border language contact. Zakarya Ali et al. (2026) identified relationships between cross-border mobility and lexical borrowing in Malay-speaking communities. The present findings suggest that returned migrant workers may continue to use lexical resources associated with multilingual workplaces after returning to their communities.

4.6. Shared Migration Experience and Solidarity

The first motivation identified in the study concerns solidarity among individuals who share employment experiences in Malaysia. According to the interview accounts, participants sometimes used Malaysian Malay expressions when communicating with people who had also worked abroad. They associated this practice

with familiarity, shared memories, and more relaxed interaction.

The observational data from Sempadian provide a relevant context. During the Eid al-Adha household visit, the participants had previously worked in Malaysia. Their conversation included expressions such as *pusak* and *selipar*, which appeared during ordinary exchanges about household matters.

Another observed interaction involved two women who had become friends while working in Malaysia. One of their exchanges included:

(6) *Kasut baru ya?*

'New shoes, huh?'

The expression occurred in a familiar interaction between former co-workers. Together with the participants' reported preference for using Malaysian Malay with fellow returnees, this context provides support for interpreting code-mixing as a possible resource for expressing familiarity.

From Gumperz's (1982) perspective, such language choices may function as a *we-code*, signalling shared experience and group belonging. Myers-Scotton's (1993) markedness perspective also suggests that a familiar linguistic choice may be relatively unmarked within a group whose members share similar migration histories.

Nevertheless, the interpretation requires an important distinction. The participants' accounts support the association between code-mixing and solidarity, while the theoretical explanation of *we-code* represents the researchers' interpretation of that association. The available observations do not establish that every Malaysian Malay expression was deliberately selected to communicate solidarity.

This finding complements earlier studies demonstrating the role of social relationships in multilingual communication (Bhatia & Ritchie, 2013; Hasbi et al., 2025; Rijal, 2021). It also suggests that language practices acquired during migration may continue to support interpersonal relationships after return.

4.7. Displaying Migration Experience as Social Identity

The second motivation concerns the use of Malaysian Malay expressions as a way of communicating previous migration experience. During interviews, participants reported that they sometimes used Malaysian Malay when speaking with friends who had never worked in Malaysia. Their explanations associated this practice with the desire to appear experienced and *keren* ('cool').

These accounts indicate that some participants understood Malaysian Malay expressions as markers of their experiences abroad. Rather than describing their employment histories directly, they reported using language as one way of presenting themselves as individuals who had worked in another country.

This interpretation is supported primarily by participants' interview accounts rather than direct observations of interactions with non-migrant audiences. Consequently, the findings establish how participants described their own intentions but do not independently demonstrate how listeners interpreted those expressions.

Myers-Scotton's (1993) markedness perspective provides a useful explanation of this reported practice. In interactions where Malaysian Malay is not expected, its use may become socially noticeable and may communicate a distinctive personal background. Hall & Níle (2015) similarly explain that language choice can contribute to identity construction, while Nugroho (2022) documented the use of Korean expressions among returned Indonesian migrant workers in connection with overseas experience and social identity.

The Sambas findings extend these discussions by suggesting that Malaysian Malay may serve as a linguistic resource through which returned workers present their migration histories in their home communities.

However, the study cannot conclude that this practice necessarily increases social status or produces recognition among non-migrant listeners. Such conclusions would require direct evidence from audience responses. The present findings are therefore best interpreted as evidence of reported self-presentation rather than demonstrated social-status change.

4.8. Residual Habit and Lexical Accessibility

The third motivation concerns the continued use of linguistic expressions acquired during employment in Malaysia. Interview participants explained that certain Malaysian Malay words remained familiar after returning home and sometimes appeared spontaneously in their everyday speech.

This explanation differs from the motivations associated with solidarity and social identity. Whereas those motivations may involve deliberate or socially meaningful language choices, residual habit concerns

the continued availability of familiar lexical expressions.

Examples include vocabulary associated with employment, transportation, and routine interaction:

(4) *Dahulu, aku bekerja di Malaysia sering diminta overtime.*

'When I worked in Malaysia, I was often asked to work overtime.'

(14) *Saya gosan-kan mobil Bapak.*

'I'll reverse your car, Sir.'

(16) *Silakan dicuba makanannya, Ibu.*

'Please try the food, Ma'am.'

These utterances contain expressions associated with activities that may have been familiar during participants' experiences in Malaysia. The interpretation is consistent with earlier research linking prolonged language contact to continued linguistic influence among returned migrant workers (Al-Pansori, 2014; Din-iarti et al., 2017).

However, the examples alone do not establish whether the expressions were produced automatically or deliberately. The interpretation of residual habit is supported by participants' reported experiences of continuing to use familiar Malaysian Malay vocabulary.

The distinction is important because a single expression may have different meanings across communicative situations. A word used spontaneously in one conversation may be used deliberately in another. Accordingly, residual habit should not be treated as mutually exclusive from solidarity or identity-related motivations.

4.9. Integrating Linguistic Forms, Migration Experiences, and Social Motivations

The findings demonstrate that Malaysian Malay code-mixing among returned migrant workers involves both structural and social dimensions. At the linguistic level, the corpus contains lexical insertions, phrase-level constituents, and utterances involving multiple Malaysian Malay elements. Some examples also display structural characteristics that complicate the distinction between insertion and congruent lexicalization.

At the social level, the interview accounts identify three recurring motivations: solidarity, social identity, and residual linguistic habit. These motivations are associated with different communicative circumstances, but they should not be understood as fixed explanations for particular grammatical forms.

The central contribution of the study is the integration of linguistic classification with migration-related social interpretation. Muysken's (2000) typology helps explain how Malaysian Malay elements appear within Indonesian utterances. Gumperz's (1982) perspective provides an interpretation of group solidarity, while Myers-Scotton's (1993) framework helps explain how language choices may communicate social positioning.

The relationship between these dimensions is interpretive rather than causal. For example, *kasut* may appear in a familiar conversation between former co-workers, but its grammatical form does not establish the speaker's intention. Similarly, a Malaysian Malay expression used in front of non-migrant friends cannot automatically be interpreted as a claim to higher social status.

The findings also highlight the importance of distinguishing reported motivations from observed interactional evidence. The strongest observational support concerns the presence of Malaysian Malay elements in familiar conversations involving returned workers. Interpretations concerning deliberate status display and residual habit rely more heavily on participants' interview accounts.

These findings suggest that migration-related language contact should not be examined solely through linguistic interference. The continued use of Malaysian Malay may also reflect the social relationships and personal experiences associated with working abroad. Nevertheless, the available corpus does not establish how frequently these motivations occur or whether they are recognised in the same way by other members of the community. Table 2 summarises the three motivations, their supporting evidence, and the limits of their interpretation.

Table 2. Evidence-Based Interpretation of Social Motivations for Code-Mixing

Dimension	Solidarity	Social identity/status display	Residual habit
Participants' reported explanation	Using Malaysian Malay with fellow returnees creates familiarity and recalls shared experiences	Using Malaysian Malay can communicate previous employment abroad and make speakers appear experienced	Familiar Malaysian Malay expressions continue to be used after returning home

Relevant empirical material	Observed conversations among former migrant workers in Sem-padian	Interview accounts describing communication with non-migrant friends	Interview accounts of continued lexical familiarity
Illustrative utterances	(2), (3), and (6)	No specific utterance independently establishes audience recognition	(4), (14), and (16)
Theoretical interpretation	Gumperz's <i>we-code</i> and familiar or unmarked language choice	Myers-Scotton's markedness perspective and identity construction	Continued lexical familiarity following prolonged language contact
Evidence limitation	Observations do not establish deliberate solidarity in every instance	Non-migrant audience responses were not directly examined	Spontaneity cannot be established from interview-derived examples alone
Interpretive conclusion	Code-mixing may support familiarity and shared group identity	Code-mixing may function as reported self-presentation	Code-mixing may reflect continued language habits

Taken together, the findings suggest that the linguistic consequences of migration extend beyond the period of employment abroad. Malaysian Malay elements remain present in the linguistic repertoires of returned workers and may acquire different social meanings depending on the communicative context. At the same time, the close relationship between Indonesian and Malaysian Malay creates structural overlap and potential semantic ambiguity. These characteristics make the Sambas case relevant to broader discussions of language contact, migration, and identity, while also demonstrating the importance of cautious interpretation when the available evidence is limited.

5. CONCLUSION, IMPLICATION, SUGGESTION, AND LIMITATIONS

This study examined the forms of Malaysian Malay code-mixing in the Indonesian speech of returned migrant workers from Sambas Regency and explored the social and linguistic factors associated with its use. The analysis of 25 utterances identified lexical insertions, phrase-level constituents, and expressions containing multiple Malaysian Malay elements. Although insertion emerged as a prominent structural pattern, several utterances showed overlapping grammatical characteristics that complicated the distinction between insertion and congruent lexicalization. The findings also revealed the continued use of vocabulary associated with everyday activities in Malaysia, morphological integration between closely related varieties, English-derived lexical elements, and words with divergent meanings in Indonesian and Malaysian Malay.

The findings further identified three recurring motivations associated with code-mixing: solidarity, social identity, and residual linguistic habit. Participants reported using Malaysian Malay expressions to maintain familiarity with fellow returnees, communicate their experiences of working abroad, and continue linguistic practices acquired during migration. Observational data provided contextual evidence of Malaysian Malay expressions appearing in conversations among individuals with shared migration experiences. However, interpretations concerning deliberate status display and spontaneous lexical retrieval relied more heavily on participants' interview accounts. These motivations should therefore be understood as reported or contextually interpreted explanations rather than independently established causes of particular linguistic forms.

The study suggests that Malaysian Malay code-mixing among returned migrant workers is not adequately understood solely as linguistic interference. It may also reflect the continuing influence of migration on everyday communication and the ways speakers relate their overseas experiences to social life at home. At the same time, the structural similarities between Indonesian and Malaysian Malay create potential semantic ambiguity, particularly when similar lexical forms carry different meanings. The findings therefore highlight the importance of examining linguistic structure and social interpretation together, while recognising that the social meanings of code-mixing depend on the available interactional evidence.

This study makes a context-specific contribution to research on language contact, migration, and code-mixing by connecting the structural characteristics of Malaysian Malay elements with the social experiences reported by returned migrant workers. Previous studies of Malay-Indonesian contact among migrant workers have primarily examined linguistic interference and the continued influence of language exposure after migration (Al-Pansori, 2014; Diniarti et al., 2017). The present study complements this perspective by considering how the same linguistic resources may also be associated with shared migration experiences, identity-related self-presentation, and continued language habits.

The analysis also demonstrates the value of combining complementary sociolinguistic perspectives. Muysken's (2000) typology provides a basis for examining insertion and possible structural overlap between closely related varieties. Gumperz's (1982) *we-code* concept offers an interpretation of how linguistic choices may communicate shared experience and solidarity, while Myers-Scotton's (1993) markedness perspective helps explain how speakers may use language to negotiate social positioning. These perspectives are applied as complementary analytical tools rather than as evidence of a fixed causal relationship between linguistic forms and social motivations.

A further contribution concerns the identification of English-derived vocabulary and lexical items with divergent meanings across Indonesian and Malaysian Malay. These features illustrate how migration-related language contact may involve multiple linguistic influences and potential semantic ambiguity. Nevertheless, given the limited corpus of 25 utterances, the findings should be regarded as exploratory and context-specific. The study does not establish a new general theory of code-mixing but demonstrates the analytical usefulness of examining linguistic forms, migration experiences, and reported social motivations within an integrated framework.

The findings offer preliminary considerations for language education and communication practices in migrant-sending communities such as Sambas Regency. For Indonesian language teachers, the documented expressions may provide locally relevant examples for discussing language variation, code-mixing, and differences in lexical meaning between Indonesian and Malaysian Malay. Particular attention may be given to expressions such as *percuma*, *senang*, *pusing*, *bandar*, *kereta*, and *budak*, whose meanings may differ across the two varieties. Such examples could support students' awareness of linguistic diversity without presenting the language practices of returned migrant workers simply as errors.

The findings may also inform communication awareness among local government agencies, community organisations, and service providers working with returned migrant workers. Recognising that familiar-looking words can carry different meanings may be useful when preparing public information or communicating across groups with different migration experiences. However, these implications represent potential applications of the findings rather than interventions evaluated by the study. The effectiveness of educational materials or communication strategies based on these observations would require further development and assessment.

Several limitations should be considered when interpreting the findings. First, the analysis was based on a relatively small corpus of 25 code-mixed utterances, comprising seven naturally observed utterances and eighteen interview-derived examples. The predominance of interview-derived data limits the extent to which the findings can represent spontaneous everyday language use. Interview examples may reflect expressions recalled or reported by participants rather than their actual frequency or distribution in naturally occurring conversations. Consequently, the study cannot determine how commonly particular forms occur among returned migrant workers or whether insertion is the dominant pattern in the wider population.

Second, the interpretation of social motivations relied partly on participants' explanations of their own language practices. Although observational data provided contextual support for some interpretations, the study did not systematically include non-migrant listeners or examine their responses to Malaysian Malay expressions. Therefore, claims concerning social recognition, perceived status, and misunderstanding remain provisional. The analysis identifies participants' reported intentions and potential semantic ambiguity but does not establish how these expressions are consistently interpreted by other community members.

Third, the linguistic boundaries between Indonesian, Sambas Malay, Sarawak Malay, and other local varieties were not examined comprehensively. Because these varieties share historical and structural characteristics, some expressions may be difficult to attribute exclusively to one variety. This limitation is particularly relevant to the distinction between insertion and congruent lexicalization. The study was also restricted to three research locations in Sambas Regency, limiting the extent to which its findings can be transferred to other migrant-sending communities.

Future research should address these limitations through larger collections of naturally occurring conversations involving returned migrant workers. Such data would allow researchers to examine the distribution of linguistic forms more systematically and assess whether particular expressions vary according to interlocutor, communicative setting, or time since returning from Malaysia. Longitudinal research could further investigate whether the use of Malaysian Malay changes as workers readjust to their home communities.

Additional studies should include non-migrant listeners to examine how Malaysian Malay expressions are understood and whether speakers' reported intentions correspond with audience interpretations. Comparative research involving Sambas Malay, Sarawak Malay, and Indonesian could also provide a stronger linguistic basis for distinguishing insertion, congruent lexicalization, and shared lexical forms. Finally, examining language use among workers while still employed in Malaysia, as well as their communication through social media after returning home, could provide a broader understanding of how migration experiences shape linguistic practices across different settings.

These directions would help determine whether the patterns identified in this exploratory study are also present in larger and more varied communicative contexts, while strengthening the empirical basis for understanding the relationship between language contact, migration, and social identity.

REFERENCES

- Al-Pansori, M. J. (2014) Interferensi bahasa Melayu Malaysia terhadap bahasa Indonesia pada masyarakat TKI Kecamatan Sikur. *Mabasan*, 8(2), 141–149. <https://doi.org/10.62107/mab.v8i2.92>.
- Amin, M. F., & Suyanto, S. (2017) Pergeseran dan pemertahanan bahasa ibu dalam ranah rumah tangga migran di Kota Semarang. *Nusa: Jurnal Ilmu Bahasa dan Sastra*, 12(1), 15–26. <https://doi.org/10.14710/nusa.12.1.15-26>.
- Bhatia, T. K., & Ritchie, W. C. (Eds.) (2013) *The Handbook of Bilingualism and Multilingualism* (2nd ed.). Wiley-Blackwell.
- Dilah, T., Chong, S., & Muhammad Jazlan, A. K. (2023) Language attitude among the border community at Telok Melano, Malaysia and Temajuk, Indonesia: A preliminary study. *Suvannabhumi: Multi-disciplinary Journal of Southeast Asian Studies*, 15(1), 229–254.
- Diniarti, D. A., Amrulloh, A., & Syahdan, S. (2017) Interferensi bahasa Melayu terhadap bahasa Indonesia pada TKI di Dusun Getap Kecamatan Suralaga. *Prasasti: Journal of Linguistics*, 2(2), 205–221. <https://doi.org/10.20961/prasasti.v2i2.1986>.
- Gumperz, J. J. (1982) *Discourse Strategies*. Cambridge University Press.
- Haerudin, H., & Diniarti, D. A. (2021) Campur kode di lingkungan masyarakat Pancor Sanggeng, Kabupaten Lombok Timur. *Berajah Journal*, 1(1), 50–55. <https://doi.org/10.47353/bj.v1i1.19>.
- Hall, K., & Nilep, C. (2015) Code-switching, identity, and globalization. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The Handbook of Discourse Analysis* (2nd ed., pp. 597–619). Wiley Blackwell.
- Hasbi, A. I. P., Abduh, A., Basri, M. A., & Tahir, M. (2025) The use of code mixing by foreign refugees in multilingual environment of Makassar City towards English language teaching in Indonesia. *Klasikal: Journal of Education, Language Teaching and Science*, 7(1), 229–239. <https://doi.org/10.52208/klasikal.v7i1.1283>.
- Herawati (2016) Campur kode dalam peristiwa komunikasi di lingkungan sekolah SMA Negeri 1 Kabangka. *Jurnal Humanika*, 16(1).
- Ladegaard, H. J. (2018) Codeswitching and emotional alignment: Talking about abuse in domestic migrant-worker returnee narratives. *Language in Society*, 47(5), 693–714. <https://doi.org/10.1017/S0047404518000933>.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014) *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.). Sage.
- Muysken, P. (2000) *Bilingual Speech: A Typology of Code-Mixing*. Cambridge University Press.
- Myers-Scotton, C. (1993) *Social Motivations for Codeswitching: Evidence from Africa*. Clarendon Press.
- Nugroho, S. A. (2022) Code mixing used by returned migrant workers in their social media posts. *INContext: Studies in Translation and Interculturalism*, 2(3), 105–126. <https://doi.org/10.54754/incontext.v2i3.33>.
- Republic of Indonesia (2017) *Law No. 18 of 2017 on the Protection of Indonesian Migrant Workers*.
- Rijal, A. S. (2021) Penggunaan bahasa dalam ranah pariwisata: Studi di kawasan Taman Nasional Bantimurung Bulusaraung Maros, Sulawesi Selatan. *Jurnal Sasdaya*, 5(1), 37–51.
- Samatan, N., Yusriyah, K., Sariyati, Nurainy, R., & Robingah (2022) Code mixing among millennials of Banggai ethnic. *International Journal of Management Studies and Social Science Research*, 4(5), 94–108. <https://doi.org/10.56293/IJMSSSR.2022.4512>.
- Sarmin, S., Karim, A., & Nur, Y. (2023) Campur kode bahasa Melayu Malaysia terhadap penggunaan bahasa Indonesia di wilayah perbatasan Pulau Sebatik Kalimantan Utara. *Jurnalistrendi: Jurnal Linguistik, Sastra, dan Pendidikan*, 8(1), 31–41. <https://doi.org/10.51673/jurnalistrendi.v8i1.1351>.

Sudaryanto (1993) *Metode dan Aneka Teknik Analisis Bahasa*. Duta Wacana University Press.

Weinreich, U. (1953) *Languages in Contact: Findings and Problems*. Linguistic Circle of New York.

Zakarya Ali, H., Noordin, A. G., & Putra, A. D. (2026) Lexical borrowing and morphosyntactic variation in cross-border Malay speech communities. *Jurnal Skripta*, 11(2), 146–160.
<https://doi.org/10.31316/skripta.v11i2.9727>.