International Journal of Language and Teaching (IJOLATE)
E-ISSN: XXXX-XXXX (Online)/P-ISSN: XXXX-XXXX
VOL. 1, Number 2 (2025)

Beyond the Boundaries: Multilingual Code-Mixing in WhatsApp Conversations Among EFL University Students

Bunga Istiqomah
bungaistiqomah146@gmail.com
InstitutAgama Islam Negeri Curup, Indonesia

Dino Anres Putra
dinoanresputra@gmail.com
InstitutAgama Islam Negeri Curup, Indonesia

Alya Faaizah
alyafaaizah@gmail.com
InstitutAgama Islam Negeri Curup, Indonesia

Abstract: This study investigates the types and functions of code-mixing in the informal WhatsApp conversations of EFL university students at the English Tadris Study Programme (TBI 4C), IAIN Bengkulu. Using a qualitative descriptive approach, 34 naturally occurring utterances were collected from seven Semester 4 students through non-participant observation over one month (April 12 to May 12, 2026). The data were analyzed using Muysken's (2000) typological framework, which classifies code-mixing into Insertion, Alternation, and Congruent Lexicalization, and Hoffmann's (1991) functional framework, which identifies the communicative purposes behind code-mixing. The findings show that Insertion was the dominant type (85.3%), followed by Congruent Lexicalization (8.8%) and Alternation (5.9%), with students mostly embedding English words into an Indonesian or regional language base. In terms of function, lexical gap filling was the most frequent (50.0%), followed by affective and emphasis function (35.3%), group solidarity (11.8%), and prestige and affiliation (2.9%). These results suggest that code-mixing among these students is not a sign of linguistic deficiency, but a purposeful and sophisticated communicative strategy that reflects their trilingual identity across Indonesian, Bengkulu dialect, and English in everyday digital communication.

Keywords: Code-Mixing, EFL Students, Multilingual

Introduction

Think about how you talk with your friends on WhatsApp. Chances are, you do not stick to just one language the whole time. This is actually a very common phenomenon called code-mixing, which happens when speakers draw on two or more languages within a single conversation or even a single sentence. For EFL (English as a Foreign Language) university students in Indonesia, this is especially familiar territory, since these learners naturally move between Indonesian, regional languages, and English in their everyday lives. And with platforms like WhatsApp now central to how students communicate, these multilingual habits have found a very comfortable home in informal digital spaces, far away from the structured rules of the classroom (Ndabihonge, 2023).

The problem is, code-mixing does not always get the credit it deserves. It is often dismissed as a sign of poor language skills or incomplete learning, as if mixing languages means you cannot fully master any of them. But sociolinguistic research actually paints a very different picture. A growing number of scholars argue that code-mixing is a sophisticated communicative strategy, one that reflects multilingual identity and the ability to navigate multiple languages flexibly and

purposefully (Setyawati et al., 2023). Understanding why and how EFL students code-mix in informal digital settings can tell us a lot about the relationship between language learning, identity, and communication in today's multilingual world.

Code-mixing in WhatsApp conversations among EFL university students presents a rich and naturally occurring sociolinguistic phenomenon. These naturally occurring utterances reveal that code-mixing in this digital context is neither random nor a sign of linguistic deficiency. Rather, it reflects a sophisticated and purposeful deployment of multilingual resources to serve a range of communicative needs — from filling lexical gaps to expressing emotion and reinforcing in-group solidarity.

This study focuses specifically on the WhatsApp conversations of EFL university students at the English Tadris Study Programme (TBI 4C) at IAIN Bengkulu. The participants are seven active Semester 4 students (coded P1–P7) with multilingual backgrounds encompassing Indonesian, at least one regional language (Bengkulu dialect), and English. Data were collected through non-participant observation of both group and private WhatsApp chats over a one-month period (April 12 – May 12, 2026), yielding a total of 34 utterances containing instances of code-mixing.

The analysis of these utterances draws on two established theoretical frameworks: *Muysken's (2000)* typological framework, which distinguishes three structural types of code-mixing: the Insertion, Alternation, and Congruent Lexicalization. And *Hoffmann's (1991)* functional framework, which identifies the communicative purposes served by code-mixing, including lexical gap filling, group solidarity, prestige/affiliation, and emphasis/affective function. Together, these frameworks provide a comprehensive lens for examining both the structure and the communicative purpose of code-mixing as it emerges in this particular informal digital context.

Previous researchers have conducted studies on code-mixing in WhatsApp and other digital communication contexts. From an international multilingual perspective, Pérez-Sabater (2022) examined Catalan-English-Spanish mixing on WhatsApp, Gonzales & Tsang (2023) and Gonzales (2025) investigated Cantonese-English code-switching in Hong Kong's digital landscape, Ndabihonge (2023) analyzed multilingual WhatsApp interactions in the Burundian context, Adisaturrahimi & Hadi (2023) explored code-mixing among Pakistani bilingual immigrants on WhatsApp, and Shubhra & Krishna (2024) approached WhatsApp youth interactions from the angle of language attrition. From the perspective of Indonesian WhatsApp communication, Setiawati & Farahsani (2021), Haryati & Prayuana (2020), Vianingrum et al. (2022), Yohana (2023), and Setyawati et al. (2023) have documented code-mixing and code-switching patterns in various WhatsApp group settings among lecturers, students, and organizational members, while Husnah & Zaini (2022) and Rufaidah et al. (2021) have specifically examined Arabic-Indonesian mixing in Islamic university WhatsApp communities. More closely related to the present study, Menggo et al. (2023) analyzed insertion functions in WhatsApp code-mixing among university students, and Agung et al. (2022) examined code-mixing in student-lecturer WhatsApp interactions in an Indonesian EFL setting. However, none of these studies have specifically examined code-mixing in the WhatsApp conversations of EFL university students in a tri-lingual regional context involving Indonesian, Bengkulu dialect, and English simultaneously, nor have they applied both Muysken's (2000) typological and Hoffmann's (1991) functional frameworks together within such a context. Therefore, this study aims to fill that gap.

Beyond the gap in existing research, there is also a bigger conceptual reason why code-mixing deserves more serious attention. Studies like Menggo et al. (2023) have shown that code-mixing is not a deficit or a shortcut. It is actually a purposeful expression of multilingual competence, where speakers strategically draw on everything in their linguistic toolkit to communicate as effectively as possible. That reframing matters, and this study builds on it by documenting, classifying, and analyzing how code-mixing naturally shows up in the everyday WhatsApp conversations of TBI 4C students.

This study, titled *"Beyond Boundaries: Multilingual Code-Mixing in WhatsApp Conversations Among EFL University Students,"* takes a qualitative descriptive approach (Creswell, 2014) to look closely at real, naturally occurring language data from TBI 4C students' WhatsApp chats. It is guided by two straightforward but important questions:

1. What types of code-mixing are used by EFL university students in their informal WhatsApp conversations?
2. What functions does code-mixing serve in those conversations?

Methods

Research Design

This study employs a qualitative descriptive approach using non-participant observation as the primary method of data collection. Rather than intervening or prompting participants to produce language, this research captures code-mixing as it occurs naturally and spontaneously in participants' everyday WhatsApp conversations. This design is appropriate because the goal is not to test or measure, but to describe and interpret real linguistic behavior in its authentic context (Creswell, 2014).

Qualitative research is particularly suitable for linguistic studies of this nature, as it allows the researcher to explore the complexity of multilingual communication without reducing it to numerical variables. By observing naturally occurring language data, this study seeks to reveal the types and functions of code-mixing that emerge organically among EFL university students in informal digital communication.

Data Source & Participants

The data for this study were obtained from a WhatsApp group and private chat conversations involving EFL university students at the English Tadris Study Programme (TBI 4C). Participants were selected through purposive sampling based on specific criteria relevant to the research questions.

The participants consist of seven (7) active EFL university students (Semester 4) who: (1) have a multilingual background involving Indonesian, at least one regional language (Bengkulu dialect), and English; (2) regularly communicate via WhatsApp with their fellow students; and (3) were willing to voluntarily share their chat data for research purposes. Participants are coded as P1-P7 to ensure anonymity throughout the study.

Purposive sampling was employed because this study specifically requires participants whose linguistic repertoire includes multiple languages and who actively use WhatsApp for informal communication. This sampling strategy ensures that the data collected are rich in instances of code-mixing and directly relevant to the research questions (Merriam, 2009).

Data Collection

Research Instrument

The primary instrument used in this study is a WhatsApp Chat Documentation Sheet, which serves as a structured observation tool for recording and categorizing naturally occurring code-mixing utterances. The sheet was designed to capture: the date and participant code, the original utterance as written in the chat, the languages involved, the type of code-mixing based on Muysken's (2000) typology, the function of code-mixing based on Hoffmann's (1991) framework, and contextual notes regarding the communicative situation.

Observation Procedure

Data collection was conducted through non-participant observation over a period of one month, from April 12 to May 12, 2026. The researcher observed WhatsApp conversations asynchronously and documented utterances containing code-mixing without participating in or

influencing the conversations. The observation was carried out within the TBI 4C class setting of the English Tadris Study Programme.

Table 1. Observation Procedure

Date	12 April – 12 May 2026
Time	24-hour (Ongoing/Asynchronous)
Location	English Tadris Study Programme
Class/Setting	TBI 4C
Observer	Bunga Istiqomah
Number of Participants Observed	7 Participants (P1-P7)
Type of Activity	Any WhatsApp Conversation

Observation Protocol

To ensure authenticity and ethical compliance, the following protocols apply:

Table 2. Observation Protocol.

PROTOCOL	DETAIL
Consent	Written or verbal informed consent was obtained from all 7 participants.
Anonymization	All names, numbers, and identifying information were replaced with codes (P1–P7).
Timeframe	12 April – 12 May 2026 (1 month)
Chat Type	Any WhatsApp Conversation among EFL Peers.
Language Criteria	Only chats containing instances of code-mixing are included.

Observation Sheet (per utterance)

Each utterance containing code-mixing was recorded in the observation sheet below. A total of 34 utterances were documented from 7 participants (P1-P7) over the one-month observation period.

Table 3. Observation Sheet.

No	P	Original Utterance	Languages Used	Type (Muysken)	Function (Hoffmann)	Context Notes
1	P1	*"woi, masuk kamu ict ni?"*	ID + EN (ict)	Insertion	Lexical gap	Asking about the ICT class attendance
2	P1	*"harus prepare"*	ID + EN	Insertion	Prestige/affiliation	Responding to the discussion about filler
3	P1	*"la diancam dk lulus amb tu secara implisit"*	ID + Bengkulu + EN (implisit)	Insertion	Emphasis/affective	Reacting to implicit academic threat
4	P1	*"dk sesuai profile picture"*	ID + EN (profile picture)	Insertion	Lexical gap	Commenting on the profile picture
5	P1	*"pfp mak" fb"*	EN (pfp, fb) + ID	Congruent Lexicalization	Group solidarity	Joking about profile picture style
6	P1	*"carikan pp cak kau yen"*	ID + EN (pp) + Bengkulu	Insertion	Group solidarity	Asking Yeni to find a profile picture

7	P2	*"tula yen, besok2 ndak filler bibir"*	Bengkulu + EN (filler)	Insertion	Lexical gap	Discussing lip filler plans
8	P2	*"rekomen dk pke heels ksano?"*	ID + EN (heels)	Insertion	Lexical gap	Asking for an outfit recommendation
9	P2	*"otw tengok gen"*	EN (otw, gen) + ID	Congruent Lexicalization	Group solidarity	Casual update on activity
10	P2	*"gek benekkan am disano bung. Flat idupko, dkd challenge nyo"*	Bengkulu + EN (flat, challenge)	Insertion	Affective/emphasis	Joking about life being uneventful
11	P3	*"sp yg buat ict?"*	ID + EN (ict)	Insertion	Lexical gap	Asking who made the ICT assignment
12	P3	*"amb cop nyari footage buat konten nn ko"*	ID + EN (footage, konten)	Insertion	Lexical gap	Discussing content creation plans
13	P3	*"abis slot"*	ID + EN (slot)	Insertion	Lexical gap	Informing that slots are full
14	P4	*"Airdrop aj bsk"*	EN (Airdrop) + ID	Insertion	Lexical gap	Suggesting file sharing via Airdrop
15	P5	*"Gek saling review udah listening yu"*	ID + EN (review, listening)	Insertion	Lexical gap	Suggesting peer review of the listening task
16	P5	*"Percaya lah bes susah track ny"*	ID + EN (track)	Insertion	Affective/emphasis	Expressing difficulty with a trial
17	P5	*"Nah yg nak buat tgas mr ruly besok tulah, jdi rabu healing"*	ID + EN (healing)	Insertion	Affective/emphasis	Planning tasks and relaxation
18	P5	*"Dk bisa di download pun"*	ID + EN (download)	Insertion	Lexical gap	Reporting file download issue
19	P6	*"Jamber ict ni?"*	Bengkulu + EN (ict)	Insertion	Lexical gap	Asking about the ICT class
20	P6	*"omaygattttt tamkuttt"*	EN (oh my god) + ID	Alternation	Affective/emphasis	Expressing shock/excitement
21	P6	*"caknyo fresh"*	Bengkulu + EN (fresh)	Insertion	Affective/emphasis	Commenting on someone's appearance
22	P6	*"makeuo nian berarti today"*	Bengkulu + EN (makeup, today)	Congruent Lexicalization	Group solidarity	Commenting on the makeup worn that day
23	P6	*"harusnyo tersirat yg definition clues ti"*	ID + EN (definition clues)	Insertion	Lexical gap	Discussing reading comprehension material

24	P6	*"jawba siko yo al, caknyo belum ke makna implicit..."*	Bengkulu + ID + EN (implicit/explicit)	Alternation	Lexical gap	Explaining implicit vs explicit meaning
25	P6	*"la masuk subjunctive wish"*	ID + EN (subjunctive wish)	Insertion	Lexical gap	Notifying about the grammar topic in class
26	P6	*"kalo dasalah pas materi who, whose segalo macam tu"*	Bengkulu + EN (who, whose)	Insertion	Lexical gap	Discussing relative pronoun material
27	P6	*"idak tethingking aku"*	Bengkulu + EN (thinking)	Insertion	Affective/emphasis	Expressing disbelief
28	P6	*"Berarti skill u emang semantap itu bung #spicles"*	ID + EN (skill)	Insertion	Group solidarity	Complimenting someone's ability
29	P6	*"inini yang dibuat nambah yg di wawancara samo blue print research questionnyo bae?"*	ID + EN (blueprint, research question)	Insertion	Lexical gap	Clarifying mini research task
30	P6	*"atau full ngikut form gpt mr tu?"*	ID + EN (form, gpt)	Insertion	Lexical gap	Asking whether to follow GPT format
31	P7	*"meet di sana yaaa"*	EN (meet) + ID	Insertion	Lexical gap	Arranging a meetup
32	P7	*"apo anjir? palm cheese?"*	Bengkulu + EN (palm cheese)	Insertion	Affective/emphasis	Reacting with surprise to the mention of food
33	P7	*"malash ah ezone gek di tangkap"*	ID + EN (ezone)	Insertion	Affective/emphasis	Expressing reluctance about going to the e-zone
34	P7	*"mood amb lari al...."*	EN (mood) + ID	Insertion	Affective/emphasis	Expressing emotional impulse

Data Analysis

Data analysis in this study follows an inductive qualitative approach guided by two theoretical frameworks. The analysis is organized according to the two research questions and proceeds through several systematic steps.

Analytical Framework

To address RQ1 (types of code-mixing), the study applies *Muysken's (2000) typology of code-mixing*, which distinguishes three types: Insertion, Alternation, and Congruent Lexicalization. To address RQ2 (functions of code-mixing), the study draws on *Hoffmann's (1991) functional framework*, which identifies the communicative purposes served by code-mixing in interaction.

Table 4. Analytical Framework for RQ1.

Type	Definition	What To Look For
Insertion	Embedding of a lexical item or phrase from one language into the structure of another (Muysken, 2000).	Single words/phrases from another language are dropped into a sentence.
Alternation	A switch between two languages at clause or sentence boundaries (Muysken, 2000).	Full clause or sentence shifts to another language
Congruent Lexicalization	Two languages share the same grammatical structure, and mixed vocabulary fills the frame (Muysken, 2000).	Mixed vocabulary within a shared grammatical frame

Table 5. Analytical Framework for RQ2.

Function	Definition	What to Look For in Context
Lexical gap filling	Using an L2 word when no suitable equivalent exists in L1 (Hoffmann, 1991).	No natural equivalent in L1
Group solidarity	Code-mixing to signal shared identity or closeness within an in-group (Hoffmann, 1991).	Signals closeness, casualness, shared identity
Prestige/Affiliation	Mixing to signal education, social status, or cultural identity (Hoffmann, 1991).	Signal education, humor, or cultural meaning
Emphasis/Affection	Switching languages to intensify emotional expression or tone (Hoffmann, 1991).	Intensifies emotion or tone
Topic Shift	Language changes as the subject of conversation shifts (Hoffmann, 1991).	Language changes as the subject of conversation shifts

Analysis Procedure

The data analysis followed an iterative, multi-step procedure as described below:

1. Data Familiarization: The researcher read through all documented utterances multiple times to develop a holistic understanding of the data before applying any coding categories.
2. Identification of Code-Mixing Instances: Each utterance was examined to identify words, phrases, or clauses that represent a switch from the base language (Indonesian or regional language) to another language (typically English).
3. Classification by Type (RQ1): Identified instances were classified into Muysken's (2000) three types based on the structural position and nature of the mixed element within the utterance.
4. Classification by Function (RQ2): Each instance was then analyzed for its communicative function based on Hoffmann's (1991) framework, with reference to the contextual notes captured during observation.
5. Frequency Tabulation: The frequency of each type and function was calculated to identify dominant patterns across participants and utterances.
6. Interpretation and Discussion: The classified data were interpreted in relation to the research questions, with selected utterances presented as exemplars to illustrate key findings.

Results and Discussion

Results

Types of Code-Mixing Used by EFL University Students (RQ1)

The first research question asks: What types of code-mixing are used by EFL university students in their informal WhatsApp conversations? Drawing on Muysken's (2000) typology, each of the 34 utterances was classified into one of three types: Insertion, Alternation, or Congruent Lexicalization. The frequency distribution is presented in the table below.

Table 6. Frequency of Code-Mixing Type (N=34).

Type of Code-Mixing	Frequency	Percentage	Example Utterances
Insertion	29	85.3%	*"harus prepare, abis slot, Airdrop aj bsk"*
Alternation	2	5.9%	*"omaygattttt tamkuttt, jawba siko yo al, caknyo belum ke makna implicit..."*
Congruent Lexicalization	3	8.8%	*"pfp mak fb, otw tengok gen, makeuo nian berarti today"*
Total	**34**	**100%**	

The findings reveal that Insertion is by far the dominant type of code-mixing, accounting for 29 out of 34 utterances (85.3%). Alternation appeared in only 2 utterances (5.9%), while Congruent Lexicalization was identified in 3 utterances (8.8%). Each type is discussed in detail below with illustrative examples from the data.

Insertion

Insertion, defined as the embedding of a lexical item or phrase from one language into the grammatical structure of another (Muysken, 2000), is the most prevalent type found in this study. In the majority of cases, English words or short phrases were inserted into a base structure of Indonesian (Bahasa Indonesia) or a regional language (Bengkulu dialect).

Table 7. Table of Insertions.

No	Participant	Utterance	Languages	Type	Function	Context
2	P1	*"harus prepare"*	ID + EN	Insertion	Prestige/affiliation	Responding to the discussion about filler
8	P2	*"rekomen dk pke heels ksano?"*	ID + EN	Insertion	Lexical gap	Asking for an outfit recommendation
13	P3	*"abis slot"*	ID + EN	Insertion	Lexical gap	Informing that slots are full
18	P5	*"Dk bisa di download pun"*	ID + EN	Insertion	Lexical gap	Reporting download issue

In the utterance "harus prepare" (Data 2), the English verb prepare is inserted into an Indonesian sentence structure where the Indonesian equivalent siap or mempersiapkan would normally appear. Similarly, "rekomen dk pke heels ksano?" (Data 8) embeds the English noun heels into a Bengkulu-Indonesian sentence, likely because no single equivalent Indonesian word captures the specific footwear item as concisely. "Abis slot" (Data 13) and "Dk bisa di download pun" (Data 18) illustrate how technology-related English terms (slot, download) are seamlessly inserted into informal Indonesian speech, as these terms are well-established in Indonesian digital vocabulary.

The prevalence of insertion in this dataset aligns with Muysken's (2000) observation that insertion is the most common type of code-mixing in bilingual and multilingual communities because it requires the least structural adjustment to the base language. Speakers simply embed a foreign lexical item without altering the syntactic frame of the utterance.

Alternation

Alternation, defined as a switch between two languages at clause or sentence boundaries (Muysken, 2000), was identified in 2 utterances (5.9%). In alternation, the speaker shifts entirely from one language to another, producing a multi-clause utterance in which each segment belongs to a different language.

Table 8. Table of Alternation.

No	Participant	Utterance	Languages	Type	Function	Context
20	P6	*"omaygattt tt tamkuttt"*	EN + ID	Alternation	Affective/emphasis	Expressing shock/excitement
24	P6	*"jawba siko yo al, caknyo belum ke makna implicit..."*	Bengkulu + ID + EN	Alternation	Lexical gap	Explaining implicit vs explicit meaning

In Data 20, "omaygattttt tamkuttt", the speaker begins with an English exclamation (oh my god, stylized and elongated for expressive effect) and then immediately shifts to the Bengkulu regional language (tamkuttt, meaning 'I'm scared' or expressing fright/excitement). This represents a clear clause-boundary switch driven by emotional intensity. In Data 24, "jawba siko yo al, caknyo belum ke makna implicit...", the speaker moves from Bengkulu dialect to standard Indonesian and then inserts the English academic term implicit, producing a three-language alternation pattern that reflects the speaker's multilingual repertoire and the academic context of the conversation.

Although alternation is the least structurally integrated type of code-mixing, it is communicatively powerful as it allows speakers to exploit the pragmatic associations of each language — in this case, English for emphasis and academic discourse, and the regional language for intimacy and expressiveness.

Congruent Lexicalization

Congruent Lexicalization, defined as the use of mixed vocabulary within a shared grammatical frame in which neither language fully dominates (Muysken, 2000), was found in 3 utterances (8.8%).

Table 9. Congruent Lexicalization.

No	Participant	Utterance	Languages	Type	Function	Context
5	P1	*"pfp mak fb"*	EN + ID	Congruent Lex.	Group solidarity	Joking about profile pic style
9	P2	*"otw tengok gen"*	EN + ID	Congruent Lex.	Group solidarity	Casual activity update
22	P6	*"makeuo nian berarti today"*	Bengkulu + EN	Congruent Lex.	Group solidarity	Commenting on makeup

In Data 5, "pfp mak fb", the utterance consists entirely of abbreviated English internet slang (pfp = profile picture, fb = Facebook) integrated with the Bengkulu word mak (meaning 'like'

or an exclamation), producing a sentence where the grammatical frame is shared, and words from both languages fill the slots interchangeably. Data 9, "otw tengok gen", blends the English acronym otw (on the way) and the abbreviated word gen with the Bengkulu-Indonesian verb tengok (look/check on), in a structure where neither language is exclusively the matrix language. Data 22, "makeuo nian berarti today", similarly mixes Bengkulu (nian = very/indeed, berarti = meaning) with English (makeup, today) in a congruent frame.

These examples demonstrate that Congruent Lexicalization tends to occur in highly informal, in-group conversations where participants share the same multilingual competence and social context, allowing them to draw freely from multiple linguistic systems without requiring structural negotiation.

Looking at the 34 utterances from TBI 4C students' WhatsApp conversations, three types of code-mixing show up based on Muysken's (2000) typology, and Insertion is clearly the dominant one. A whopping 29 out of 34 utterances (85.3%) fell into this category, where students dropped English words or phrases straight into an Indonesian or regional language base. You can see this in examples like "harus prepare", "abis slot", and "Dk bisa di download pun". Congruent Lexicalization came in second with 3 utterances (8.8%), where no single language really takes over the grammatical structure, like in "pfp mak fb" and "otw tengok gen". Alternation was the rarest, showing up in just 2 utterances (5.9%), both from P6, as in "omaygattttt tamkuttt" and "jawba siko yo al, caknyo belum ke makna implicit...", where full clause-level switches between languages actually happen. The fact that Insertion dominates so heavily makes sense when you think about it: multilingual speakers in informal digital spaces tend to borrow individual words, especially English tech, academic, and lifestyle terms, without changing the grammatical structure of whatever language they are already using.

Functions of Code-Mixing in EFL University Students' WhatsApp Conversations (RQ2)

The second research question asks: What functions does code-mixing serve in EFL university students' informal WhatsApp conversations? Drawing on Hoffmann's (1991) functional framework, four functions were identified in the dataset: Lexical Gap Filling, Emphasis/Affective Function, Group Solidarity, and Prestige/Affiliation. The frequency distribution is presented in Table 10.

Table 10. Frequency of Code-Mixing Functions (N = 34).

Function	Frequency	Percentage	Example Utterances
Lexical Gap Filling	17	50.0%	*"abis slot, Dk bisa di download pun, meet di sana yaaa"*
Emphasis / Affective	12	35.3%	*"omaygattttt tamkuttt, idak tethingking aku, mood amb lari al"*
Group Solidarity	4	11.8%	*"pfp mak fb, otw tengok gen, berarti skill u emang semantap itu bung"*
Prestige / Affiliation	1	2.9%	*"harus prepare"*
Total	**34**	**100%**	

Lexical Gap Filling is the most frequently occurring function (n = 17; 50.0%), followed by Emphasis/Affective Function (n = 12; 35.3%), Group Solidarity (n = 4; 11.8%), and Prestige/Affiliation (n = 1; 2.9%). No instances of Topic Shift were identified in the dataset. Each function is discussed below.

Lexical Gap Filling

Lexical gap filling occurs when a speaker uses a word from another language because no suitable or concise equivalent exists in the base language (Hoffmann, 1991). This is the most

prevalent function in the dataset, appearing in 17 utterances (50.0%), and is closely linked to the prevalence of technology-related, academic, and lifestyle terms borrowed from English.

Table 10. Table of Lexical Gap Filling.

No	Participant	Utterance	Languages	Type	Function	Context
14	P4	*"Airdrop aj bsk"*	EN + ID	Insertion	Lexical gap	Suggesting Airdrop file sharing
7	P2	*"tula yen, besok2 ndak filler bibir"*	Bengkulu + EN	Insertion	Lexical gap	Discussing lip filler plans
25	P6	*"la masuk subjunctive wish"*	ID + EN	Insertion	Lexical gap	Notifying about the grammar topic in class

In Data 14, "Airdrop aj bsk" (Just Airdrop it tomorrow), the English brand name Airdrop is used because there is no equivalent Indonesian term for this specific file-sharing function. Similarly, in Data 7, the word filler refers to a cosmetic procedure (lip filler) that has entered Indonesian informal speech directly from English due to the absence of a natural Indonesian equivalent. Data 25 illustrates academic lexical gap filling: the English grammatical term subjunctive wish is retained because it is the established label for this grammar concept in the EFL classroom, making its replacement with an Indonesian term both awkward and potentially confusing.

The high frequency of lexical gap filling reflects the reality of multilingual EFL students whose daily lives increasingly involve digital technology, social media, beauty culture, and academic English. All domains in which English terminology is dominant, and Indonesian equivalents are either absent or rarely used in practice.

Emphasis and Affective Function

The affective function refers to code-mixing used to intensify emotional expression, convey a particular tone, or add expressive emphasis that the base language alone cannot fully achieve (Hoffmann, 1991). This function appeared in 12 utterances (35.3%), making it the second most frequent function in the dataset.

Table 12. Table of Emphasis/Affective Function.

No	Participant	Utterance	Languages	Type	Function	Context
20	P6	*"omaygattttt tamkuttt"*	EN + ID	Alternation	Affective/emphasis	Expressing shock/excitement
27	P6	*"idak tethingking aku"*	Bengkulu + EN	Insertion	Affective/emphasis	Expressing disbelief
34	P7	*"mood amb lari al...."*	EN + ID	Insertion	Affective/emphasis	Expressing emotional impulse

Data 20, "omaygattttt tamkuttt", demonstrates how the English exclamation oh my god is stretched and elongated (omaygattttt) for dramatic effect, conveying a level of exaggerated shock or excitement that the Bengkulu equivalent alone would not fully express. This reflects a common pattern in digital communication where English exclamations carry heightened emotional associations due to their prevalence in social media and popular culture. In Data 27, the morphological blend "tethingking" — where the English verb thinking is affixed with the Bengkulu prefix te- (indicating an involuntary action) — intensifies the speaker's expression of disbelief or surprise. In Data 34, "mood amb lari al", the English word mood is used in its social media sense

(meaning an overwhelming emotional urge), a nuance that the Indonesian word suasana hati would not capture in the same informal, expressive register.

These findings suggest that English functions as an emotionally expressive register for these multilingual speakers, particularly for conveying reactions that are amplified in digital and social media contexts.

Group Solidarity

Group solidarity refers to the use of code-mixing to signal shared identity, closeness, or in-group membership among interlocutors (Hoffmann, 1991). This function was identified in 4 utterances (11.8%).

Table 13. Table of Group Solidarity.

No	Participant	Utterance	Languages	Type	Function	Context
9	P2	*"otw tengok gen"*	EN + ID	Congruent Lex.	Group solidarity	Casual activity update
28	P6	*"Berarti skill u emang semantap itu bung"*	ID + EN	Insertion	Group solidarity	Complimenting someone's ability

In Data 9, *"otw tengok gen"*, the use of the English acronym otw (on the way) alongside Bengkulu-Indonesian words signals an informal, relaxed communicative register typical of close peers. The brevity and hybrid nature of the utterance itself are markers of in-group familiarity: only someone who shares the same linguistic and social repertoire would immediately understand this blend. In Data 28, *"berarti skill u emang semantap itu bung"*, the English word skill is inserted while the rest of the utterance uses informal Indonesian and the address term bung (informal for 'friend'), together creating a tone of genuine admiration and camaraderie that is characteristic of in-group peer communication.

The relatively lower frequency of this function does not diminish its significance; rather, it reflects that solidarity is often implicitly embedded in the overall code-mixing practice itself, as the very act of mixing languages signals shared linguistic competence and closeness.

Prestige and Affiliation

Prestige and affiliation refer to code-mixing as a means to signal education, cultural membership, or social status (Hoffmann, 1991). Only one instance of this function was clearly identified in the dataset.

Table 14. Table of Prestige/Affiliation Function.

No	Participant	Utterance	Languages	Type	Function	Context
2	P1	*"harus prepare"*	ID + EN	Insertion	Prestige/affiliation	Responding to the discussion about filler

In Data 2, *"harus prepare"*, the speaker uses the English verb prepare rather than the Indonesian equivalent mempersiapkan or siap. The choice of the English term in this context, which concerns a social rather than academic topic, suggests an element of prestige-driven code-mixing: the speaker signals familiarity with English and aligns themselves with a more educated or cosmopolitan identity. While this function appeared only once as the primary motivation, it likely operates as a secondary layer in several other utterances, particularly those involving academic vocabulary.

Beyond the types, the analysis also looked at why students were mixing languages in the first place, using Hoffmann's (1991) framework. Lexical gap filling topped the list at 17 utterances (50.0%), and it is not hard to see why. Students often needed to use English terms for things in tech,

beauty, or academic contexts that simply do not have a natural Indonesian equivalent, like *"Airdrop", "filler bibir",* and *"subjunctive wish".* Coming in second was emphasis and affective expression, found in 12 utterances (35.3%), where English helped students intensify emotions or reactions in a way their base language alone could not quite pull off, as in *"omaygattttt tamkuttt", "idak tethingking aku",* and *"mood amb lari al".* Group solidarity showed up in 4 utterances (11.8%), where mixing languages was really about signaling closeness and shared identity among friends, like in *"otw tengok gen"* and *"berarti skill u emang semantap itu bung*". Prestige and affiliation were the least common functions, appearing in just 1 utterance (2.9%), where the use of English hinted at an alignment with a more educated or cosmopolitan identity, as in "harus prepare". Interestingly, no instances of topic shift were found at all. Taken together, these findings make a pretty clear case that code-mixing here is intentional and communicatively driven, serving real pragmatic purposes rather than pointing to any kind of language deficiency.

Discussion

The Dominance of Insertion and Its Implications

The fact that Insertion accounts for 85.3% of all code-mixing instances in this study is perhaps the most telling finding of all. This pattern aligns closely with what Muysken (2000) describes as the default strategy in bilingual contact situations -- when one language clearly functions as the matrix, and speakers embed individual items from another language to fill the gaps. In this study, Indonesian (or the regional Bengkulu language) consistently serves as that matrix, while English supplies the embedded lexical items, especially nouns tied to technology, academia, and lifestyle.

What makes this finding more interesting is how consistently it echoes research on code-mixing in digital communication among Indonesian university students. Menggo et al. (2023) found that insertion was the dominant type in WhatsApp group chats among university students, noting that the informal and fast-paced nature of messaging encourages single-word switches rather than complex structural alternations. Similarly, Agung et al. (2022) documented a comparable pattern in student-lecturer WhatsApp interactions, where English content words were routinely inserted into Indonesian utterances. Taken together, these studies suggest that the dominance of insertion is not unique to this context but reflects a broader, stabilized pattern of English-Indonesian language contact in digital peer communication.

Lexical Gap Filling as the Primary Function

That lexical gap filling accounts for exactly half of all code-mixing instances (50.0%) points to something important about the functional role English plays in these students' communicative lives. Hoffmann (1991) observed that lexical gap filling is one of the most pragmatically driven motivations for mixing; it is not about prestige or performance, but about communicative efficiency. Speakers reach for the English term because, in practical terms, it is the most precise and readily available option.

This study confirms that observation in a very concrete way. Terms like download, Airdrop, slot, heels, filler, and subjunctive wish represent domains, technology, academic grammar, and contemporary fashion, where English remains the default register and Indonesian equivalents, even when they technically exist, are simply not part of everyday informal speech. Setyawati et al. (2023) similarly found that EFL students and lecturers code-mixed primarily for lexical reasons, particularly when discussing academic content or technology-related matters. Adisaturrahimi and Hadi (2023) made a parallel observation, noting that English words filled conceptual slots that the matrix language could not occupy as naturally or efficiently.

This pattern raises a broader pedagogical point: rather than treating these gaps as evidence of insufficient proficiency, it may be more accurate to read them as evidence of domain-specific linguistic integration. Students genuinely inhabit a multilingual academic and digital environment where English vocabulary has become functionally embedded.

Affective Code-Mixing in Digital Communication

The affective and emphasis function (35.3%) is the second most frequent in this study, and its prominence makes sense given the context. WhatsApp, as a social platform, is inherently expressive; it is where students vent, joke, encourage each other, and react emotionally. English in this context does not function merely as a source of borrowed vocabulary; it functions as an affective resource, carrying emotional weight that Indonesian equivalents often cannot replicate in the same register.

Expressions like omaygattttt, mood, and healing have been so thoroughly embedded in Indonesian digital youth culture that they function less like foreign borrowings and more like shared social codes. Perez-Sabater (2022), studying WhatsApp code-mixing across Catalan, English, and Spanish, noted that English expressions served as emotionally intensified alternatives in informal digital exchanges. Gonzales and Tsang (2023) found comparable affective motivations in Cantonese-English switching on WhatsApp in Hong Kong, reinforcing the idea that English functions as an affective amplifier in multilingual digital contexts across different regions.

Particularly striking in this study is the instance of tethingking, where the Bengkulu regional prefix te- is attached to the English root thinking. This is not mere borrowing; it is active morphological integration, evidence that these speakers are not passive recipients of English vocabulary but creative users who blend systems to achieve nuanced communicative effects. This kind of linguistic creativity is consistent with what Ndabihonge (2023) observed in multilingual WhatsApp interactions, where speakers similarly fused structural elements across languages in ways that exceeded simple insertion.

Code-Mixing as a Marker of Multilingual Identity

Stepping back from the individual functions, what the overall pattern of findings suggests is that code-mixing in these students' WhatsApp conversations is not a symptom of linguistic insecurity or incompetence. It is a sophisticated communicative practice that reflects who these students are as multilingual individuals. They navigate at least three linguistic systems simultaneously: Indonesian as the national language, a regional Bengkulu language tied to local identity, and English as the language of academic and global participation.

The co-occurrence of regional language alongside Indonesian and English is particularly worth noting. It signals that these students do not simply switch between national and global registers; they actively maintain a regional linguistic identity even as they operate within an English-medium academic environment. This kind of trilingual negotiation in everyday peer communication has also been documented by Haryati and Prayuana (2020), who found that regional and national languages co-existed with English in complex and pragmatically motivated ways in Indonesian university WhatsApp groups.

Vianingrum et al. (2022) and Setiawati and Farahsani (2021) have further shown that code-mixing in Indonesian institutional WhatsApp communication is driven not by random interference but by social functions such as solidarity, role negotiation, and community membership. The present study extends this insight to the student peer context and to a regional university setting, suggesting that code-mixing serves similar identity-constitutive functions at that level as well.

Ultimately, these findings contribute to the growing body of research on multilingualism in EFL digital contexts by reinforcing a view of code-mixing not as linguistic failure but as a resource, one that enables students to communicate more precisely, expressively, and authentically within their communities, while simultaneously negotiating their identities as regional speakers, Indonesian citizens, and emergent English users.

Conclusion

This study set out to answer two questions: what types of code-mixing do EFL university students at TBI 4C, IAIN Bengkulu, use in their WhatsApp conversations, and what communicative functions does that mixing actually serve? Looking at 34 naturally occurring utterances collected over one month, the answers turned out to be pretty consistent and revealing.

On the typology side, Insertion was the clear frontrunner, accounting for 85.3% of all instances. Students routinely dropped English words into their Indonesian or Bengkulu-dialect base, whether for a tech term like download, a grammar concept like subjunctive wish, or a lifestyle word like filler. Congruent Lexicalization came in second at 8.8%, and Alternation was the rarest at 5.9%, mostly showing up in emotionally charged moments. As for function, lexical gap filling led at 50.0%, meaning students mixed languages most often simply because the English word did the job better. Affective and emphasis expression followed at 35.3%, group solidarity at 11.8%, and prestige or affiliation at just 2.9%.

Taken together, these findings make a strong case that code-mixing among these students is not a sign of something going wrong with their language learning. It is a purposeful and sophisticated communicative strategy. These students are not stuck between languages; they are navigating three of them simultaneously, Indonesian, Bengkulu dialect, and English, and doing it with a level of pragmatic flexibility that deserves far more credit than it usually gets.

These findings carry real practical implications, particularly for EFL educators and curriculum designers. First, since code-mixing is clearly driven by communicative need rather than linguistic deficiency, teachers may want to reconsider treating it as an error to be corrected. Acknowledging students' multilingual practices in the classroom, rather than penalizing them, could actually build more confidence and engagement, especially for learners from regional language backgrounds like Bengkulu. Second, the dominance of lexical gap filling points to something worth addressing in vocabulary instruction: students are reaching for English terms in domains like technology, academic discourse, and contemporary lifestyle because their Indonesian vocabulary in those areas is simply underdeveloped. This suggests that more contextually relevant vocabulary building, grounded in the digital and social realities students actually live in, could meaningfully reduce unnecessary code-mixing while strengthening overall communicative competence. Third, for program developers, the presence of trilingual patterns involving a regional language alongside Indonesian and English is a reminder that EFL instruction in Indonesia does not happen in a linguistically neutral space. Designing materials and assessments that are sensitive to regional multilingual identities could make English learning feel less like replacing who students are and more like adding to it.

That said, this study does have its limits. The sample was small, seven participants, 34 utterances, and came from a single class at one institution over one month, so the findings cannot be generalized too broadly. The data also captured only written chat messages, which means spoken multilingual behavior and longer discourse patterns were left out of the picture.

There is a lot more to explore here. Future studies could work with larger and more diverse participant groups, bring in spoken data alongside digital communication, or trace how students' code-mixing patterns shift between academic and social contexts over time. It would also be worth examining how EFL instructors respond to or engage with students' multilingual practices, since teachers' attitudes can significantly shape how students feel about their own linguistic identities in and outside the classroom.

Ultimately, what this study hopes to leave behind is a shift in perspective: code-mixing is not interference, and it is not a shortcut. For these students, it is a resource, one that helps them communicate more precisely, express themselves more authentically, and navigate the complex multilingual world they actually live in every day for future research. A well-written conclusion leaves a strong final impression, reinforcing the study’s value and encouraging further exploration of the topic.

References

Adisaturrahimi, A., & Hadi, M. Z. P. (2023). Code mixing in WhatsApp chats made by a bilingual immigrant from Pakistan. Humanitatis: Journal of Language and Literature. https://doi.org/10.30812/humanitatis.v9i2.2612

Agung, A., Andriyani, A., Dewa, I., Devi, A., & Santika, M. (2022). Code mixing is applied in the interaction of students and lecturers on WhatsApp media. Proceedings of the First International Conference on Literature Innovation in Chinese Language, LIONG 2021, 19–20 October 2021, Purwokerto, Indonesia. https://doi.org/10.4108/eai.19-10-2021.2316554

Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). SAGE Publications.

Gonzales, W. D. (2025). Predicting language choice in a digital medium: A computational approach to analyzing WhatsApp code-switching in Hong Kong. International Journal of Bilingualism. https://doi.org/10.1177/13670069251325036

Gonzales, W. D., & Tsang, Y. (2023). The sociolinguistics of code-switching in Hong Kong's digital landscape: A mixed-methods exploration of Cantonese-English alternation patterns on WhatsApp. Journal of English and Applied Linguistics. https://doi.org/10.59588/2961-3094.1041

Haryati, H., & Prayuana, R. (2020). An analysis of code-mixing usage in WhatsApp group conversations among lecturers of Universitas Pamulang. https://doi.org/10.30605/25409190.180

Hoffmann, C. (1991). An introduction to bilingualism. Longman.

Husnah, A., & Zaini, H. (2022). Code mixing in Arabic conversation among Indonesian students of Al-Azhar University. International Journal of Arabic Language Teaching. https://doi.org/10.32332/ijalt.v4i02.5724

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE Publications.

Menggo, S., Pramesti, P. D. M. Y., & Krismayani, N. W. (2023). The insertion function in code-mixing is used on WhatsApp group chats among university students. Journal of Language Teaching and Research. https://doi.org/10.17507/jltr.1403.06

Merriam, S. B. (2009). Qualitative research: A guide to design and implementation. Jossey-Bass.

Muysken, P. (2000). Bilingual speech: A typology of code-mixing. Cambridge University Press.

Ndabihonge, A. (2023). Multilingual communication in WhatsApp interactions: Empirical analysis in the Burundian context. Journal of Humanities and Social Sciences Studies. https://doi.org/10.32996/jhsss.2023.5.6.6

Pérez-Sabater, C. (2022). Mixing Catalan, English, and Spanish on WhatsApp. Spanish in Context. https://doi.org/10.1075/sic.19033.per

Rufaidah, M. F., Salshabila, N., & Hizbullah, N. (2021). Analisis campur kode bahasa Arab dalam grup WhatsApp lembaga dakwah kampus Universitas Al Azhar Indonesia. PRASASTI: Journal of Linguistics. https://doi.org/10.20961/prasasti.v6i1.44393

Setiawati, S., & Farahsani, Y. (2021). Code-switching and code-mixing in WhatsApp group chats by FEB UMY lecturers. https://doi.org/10.2991/assehr.k.210120.146

Setyawati, A. P., Novitasari, N. F., & Lailiyah, S. (2023). Types and reasons for code-mixing in WhatsApp group communication among university EFL lecturers and students. J-ELLiT (Journal of English Language, Literature, and Teaching). https://doi.org/10.17977/um046v7i12023p1-6

Shubhra, S., & Krishna, D. (2024). An analysis of WhatsApp interactions among youth: A case of language attrition. Journal of Digital Sociohumanities. https://doi.org/10.25077/jds.1.2.63-74.2024

Vianingrum, Y., Rohmaten, L. A., & Winarti, W. (2022). Use of code mixing among circulation division members at Solopos Media Group in coordinating the events through WhatsApp

chat. Proceedings of the International Conference on Science, Health, and Technology. https://doi.org/10.47701/icohetech.v3i1.2232

Yohana, F. M. (2023). Penggunaan alih kode dan campur kode dalam percakapan grup WhatsApp Halan-Halan Syantix. Prosiding Konferensi Berbahasa Indonesia Universitas Indraprasta PGRI. https://doi.org/10.30998/kibar.27-10-2022.6297

Empty Page