

Code-Mixing between Indonesian and English in TikTok Comments on Book-Related Content: An Analysis of Forms and Communicative Functions

Verna Virginia Teten[1], Badri[2]
Universitas Terbuka[1]
Universitas Primagaraha[2]
badrikumana2@gmail.com

Riwayat Artikel
Diterima
Maret 2026
Revisi
April 2026
Terbit
Mei 2026

Keywords:
code-mixing; sociolinguistics; TikTok comments; communicative function; book content

ABSTRACT
This study aims to describe the patterns of Indonesian–English code-mixing and their communicative functions in comments on three book-related TikTok posts. A qualitative descriptive method with a sociolinguistic approach was employed. The initial dataset consisted of 45 comments gathered from three posts concerning a list of Indonesian literary works, Indonesian books that left a deep impression on readers, and the book Insecurity Is My Middle Name. Following a selection process based on criteria regarding the use of both languages within the comment structures, 34 comments were analyzed. Code-mixing patterns were classified using Muysken's typology, while communicative functions were analyzed using the framework proposed by Appel and Muysken. The results indicate that the "insertion" pattern appeared in 22 comments (64.7%) and "alternation" in 12 comments (35.3%); no instances of "congruent lexicalization" were found. The expressive function was the most dominant, appearing in 23 comments (67.6%), followed by the directive function in seven comments (20.6%), the referential function in two comments (5.9%), and the phatic and metalinguistic functions in one comment each (2.9%). The findings suggest that, within this limited corpus, code-mixing is primarily used to convey readers' assessments and reactions, as well as to ask questions or provide reading recommendations. These findings are specific to the comments from the three analyzed posts and should not be generalized to the entire community of readers or TikTok users.

©2025 Universitas Bina Bangsa
How to cite (*in APA Style*): Teten, V. V. ., & Badri. (2026). Code-Mixing between Indonesian and English in TikTok Comments on Book-Related Content: An Analysis of Forms and Communicative Functions. *Prakata: Jurnal Bahasa Dan Sastra Indonesia Serta Pembelajaran*, *3*(1), 49–66. *Retrieved from https://ejournal.lppmbinabangsa.ac.id/index.php/prakata/article/view/370*

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

PENDAHULUAN

Language is used not merely as a system of signs governed by grammatical rules, but also as a resource for building social relationships, expressing attitudes, and negotiating meaning within interactions. Language choices may shift depending on the interlocutor, the topic, the communicative purpose, and the situation in which the communication takes place. Sociolinguistics examines the relationship between language use and such social environments. Sociolinguistic textbooks in Indonesia, including those by Tricahyo (2019), Nurpadillah (2023), Susylowati et al. (2024), Yanti et al. (2024), and Nasir (2025), explore the connection between language and society, language variation, code selection, language contact, bilingualism, and the phenomena that arise when speakers employ more than one language. This framework is highly relevant for analyzing social media comments as linguistic practices occurring within specific communities, rather than viewing them simply as strings of words devoid of context.

Language contact is evident in the practice of code-mixing, the incorporation of elements from another language into speech or text constructed primarily in a specific language. In everyday communication, Indonesian may appear alongside English elements, such as individual words, phrases, clauses, or segments of sentences. The use of more than one language in communication is also linked to the process of language production and the linguistic environment of the speaker. Siagian and Damanik (2024), in a comparative study of language production in monolingual and multilingual children, reported distinct characteristics in language production between the two groups. Their research also addresses the emergence of mixed-language usage among multilingual children, albeit within a context that differs from written interactions on social media. These findings serve as a supporting reference for understanding that the use of mixed language is a phenomenon that can be examined through the lens of language production processes, while accounting for the characteristics of the speakers and the context of their communication.

The use of English elements may relate to specific terminology, a speaker's knowledge, communication habits, or linguistic styles that have emerged within digital environments. Analytically, however, not every word derived from English should automatically be categorized as code-mixing. Some words have been assimilated or have become established terms within the Indonesian language, while others retain recognizable English forms, structures, or constructions. This distinction is crucial to ensure that researchers do not conflate code-mixing with lexical borrowing, proper nouns, titles of works, or linguistic forms of ambiguous origin.

The distinction between code-mixing and code-switching also requires clarification from the outset. In this study, code-mixing is treated as an umbrella term for the use of elements from two languages within a single comment, with its structural classification following the three patterns discussed by Muysken (2000): insertion, alternation, and congruent lexicalization. Insertion occurs when elements from one language are embedded into the grammatical framework of another. Alternation refers to the switching between segments that each retain their own respective language patterns. Congruent lexicalization, meanwhile, describes a situation where lexical elements from different languages appear within a structure that shares a common grammatical framework. These three patterns, originally developed from the study of bilingual code-mixing, have been adapted here for the analysis of written comments. Consequently, their application should be viewed as an operational analytical tool, rather than an implication that digital comments are entirely identical to the spoken conversations that were the focus of the original typological framework.

The development of social media has expanded the scope for using mixed language within written communication—a form of interaction that is typically rapid, concise, and often multimodal. Comments may incorporate capital letters, repeated characters, abbreviations, emojis, quotations, or visual elements, all of which work together to construct a message. While these elements influence how a message is interpreted, they do not necessarily constitute "code-switching." For instance, capital letters can intensify an expression, emojis can provide affective cues, and abbreviations can condense an utterance. In the study of online comments, the context of the original post is crucial; a short sentence could function as a compliment, a question, a recommendation, a response, or an attempt to sustain an interaction. Consequently, analyzing language use in comments requires an examination of both the linguistic structure and the communicative context.

TikTok provides a space for users to disseminate content and interact through comments. It can also be understood as a digital environment where language is employed across various communicative and educational activities. Research by Mosende et al. (2024) on the use of educational TikTok videos for grammar instruction demonstrates that the platform has the potential to foster user engagement in language-related activities. Although their study focuses on grammar learning rather than code-switching within comments, it offers a relevant context for viewing TikTok as a medium for language-based interaction. In the present study, this context serves as a foundation for examining book-themed comments as a form of written communication on social media.

One context that has emerged on this platform involves discussions about books—including recommendations, reviews, reading experiences, and reactions to specific titles and characters. Christie et al. (2024) analyzed 131 videos tagged with #BookTokIndonesia, reporting that the content frequently focuses on specific authors and books, while also highlighting themes of community and reader identity. Although this study provides context regarding the practice of sharing reading experiences within the Indonesian BookTok community, its primary unit of analysis is the content of the videos themselves. Meanwhile, Toisuta (2024) examines the use of "netspeak" within BookTok and Bookstagram, drawing attention to the use of symbols, emoticons, abbreviations, and a more conversational style. Toisuta's findings demonstrate the relevance of a linguistic approach to understanding interactions centered on books in digital spaces, though the study does not specifically classify the patterns of Indonesian–English code-mixing present in the comments analyzed.

Research on code-switching in social media has been conducted using various approaches and data sources. Tarihoran et al. (2022) examined the relationship between social media usage and Indonesian–English code-switching among Generation Z, employing questionnaires and interviews. Rizky et al. (2024) investigated the attitudes of 500 Indonesian Twitter users toward English–Indonesian code-switching, reporting a generally positive tendency across cognitive, affective, and conative dimensions. Astuti et al. (2023) compiled and analyzed a dataset of Indonesian–English code-switching for sentiment analysis tasks on Twitter, focusing on the processing of mixed-language data using computational models. Together, these studies demonstrate that digital code-switching can be explored through the lenses of user behavior, language attitudes, and the characteristics of textual data. However, these approaches and research questions differ from a descriptive analysis of the communicative functions of code-switching within book-related comments.

Research utilizing TikTok comments as data aligns closely with the focus of this study. Mustikarani and Pratiwi (2025) analyzed comments on the TikTok account @sabrinachairunnisa_ and

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

identified the insertion of linguistic elements in the form of words, phrases, and clauses. Amiroh et al. (2025) examined code-switching and code-mixing within the content of a TikTok influencer, while Sari et al. (2026) analyzed comments from K-pop fans on TikTok, linking the use of code-mixing to self-expression, appreciation, identity, and community solidarity. Tarigan (2026) also analyzed 25 social media comments containing mixed languages, applying theories of code-mixing and communicative functions. These studies confirm that social media provides relevant data for observing language mixing and that communicative functions cannot always be explained solely through grammatical forms. At the same time, these studies demonstrate that code-mixing in social media comments is not an entirely novel phenomenon.

Based on the aforementioned prior research, the gap addressed in this article is contextual and descriptive in nature. This study does not claim that code-switching in TikTok comments or the language used within book-related communities has never been investigated. Instead, it focuses on comments found on three book-themed posts, each covering a different topic: a list of Indonesian literary works, Indonesian books that left a lasting impression on readers, and the book *Insecurity Is My Middle Name*. The data analyzed consist not of the video content or the broader identity of the community, but rather of the structure of the comments—which incorporate both Indonesian and English—and the communicative functions that can be interpreted from their specific contexts. This scope is intended to provide a focused analysis of how English linguistic elements are integrated or interspersed when users respond to books and their own reading experiences.

To analyze communicative functions, this study employs six functions associated with Appel and Muysken (1987): referential, directive, expressive, phatic, metalinguistic, and poetic. The referential function relates to the conveyance of information; the directive function concerns attempts to guide or elicit a response from an interlocutor; the expressive function highlights the speaker's attitudes or feelings; the phatic function pertains to the initiation or maintenance of social relationships; the metalinguistic function focuses on language or expressions themselves; and the poetic function emphasizes the form of the message. In digital comments, these functions may co-occur; therefore, this study identifies a single dominant function for the purpose of tabulation, while also noting any additional functions supported by the context. This functional classification is not intended to infer the personal motives or psychological states of the users.

Based on the preceding discussion, this study addresses two questions: (1) what are the patterns of Indonesian–English code-mixing found in the comments on three book-themed posts on TikTok; and (2) what are the dominant communicative functions manifested in these comments? The study aims to identify and describe the patterns of code-mixing using Muysken's (2000) typology, and to interpret their communicative functions through the framework proposed by Appel and Muysken (1987). The contribution of this research lies in providing a descriptive mapping of a specific corpus of comments, along with an account of the observed code-mixing instances that can be examined further. Given that the data selection was not random and was limited to only three posts, the findings are not intended to be representative of all TikTok users, all book readers, or the entire Indonesian "BookTok" community.

METODOLOGI PENELITIAN

This study employs a qualitative descriptive method with a sociolinguistic approach. This design was chosen because the research questions focus on describing patterns of language mixing and interpreting their communicative functions within the context of comments, rather than testing relationships between

variables or measuring cause-and-effect. Numerical data—presented as counts and percentages—serve as a descriptive summary to illustrate the distribution of the coding, while the primary analysis remains grounded in a qualitative interpretation of the comments. The use of figures in this study is not intended to provide population estimates, but rather to offer an overview of the corpus under analysis.

The analytical procedure follows a deductive qualitative content analysis approach. Elo and Kyngäs (2008) explain that content analysis can be conducted either inductively or deductively; the deductive approach is particularly useful when the category structure is derived from existing knowledge or a pre-established framework. In this study, the categories for code-mixing patterns are based on Muysken (2000), while the communicative functions are drawn from Appel and Muysken (1987). Although these initial categories were established beforehand, each data point is examined within its specific context to ensure that the coding process is not merely a mechanical search for English words.

The data source consists of comments on three book-themed TikTok posts, which have been documented by the data collector. The first post, from the account @velsjournal, discusses a list of Indonesian literary works to read in 2025. The second post, from the account @maybeitskinna, focuses on Indonesian books that have left a profound impression on readers. The third post, from the account @cakepuffe, discusses the book “Insecurity Is My Middle Name”. Based on the provided URLs, the first source is a video post, while the other two are photo/slide posts. Therefore, the terms "post" or "content" are used to refer to all the sources, rather than labeling all three as videos. The dates of these three posts are December 8, 2025, June 24, 2025, and January 18, 2024, respectively.

The unit of analysis consists of comments on TikTok. The initial dataset comprises 45 comments, with 15 comments drawn from each post. The selection of sources and comments was purposive, based on the relevance of the posts to the book-related theme and the availability of comments that could be observed and transcribed. Given that, based on the data collection experience, there was a limited number of posts containing a sufficient mix of Indonesian and English, this study ultimately utilizes three posts. This limitation is acknowledged as a boundary of the corpus, rather than serving as a basis for claiming that these three posts represent the entire BookTok community.

Inclusion criteria encompass comments that demonstrate the use of Indonesian alongside English elements within their structure. Such patterns may involve the inclusion of English words or phrases within an Indonesian construction, or the alternation between segments identifiable as English and Indonesian. Exclusion criteria cover comments written entirely in English that merely mention the title or name of an Indonesian book; comments composed in only one language; and abbreviations whose source language cannot be determined from the text—for instance, the term "ovt" found in one of the comments. Furthermore, meme formats that simply pair a book title with an English-language reaction—lacking a clear structure of language mixing—are not counted as primary data.

Of the 45 transcribed comments, 34 met the criteria for analysis, while 11 were excluded. For content C1, 12 out of 15 comments were included in the analysis; for C2, eight out of 15; and for C3, 14 out of 15. These percentages merely describe the composition of the documented sample. They do not represent the prevalence of code-switching across all comments in each post, as the data collection process did not encompass every comment, and the initial selection sequence was not recorded in detail.

Code-mixing patterns were analyzed using Muysken’s (2000) typology. The category of "insertion" is applied when an English word or

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

phrase appears as an element embedded within an Indonesian-language comment framework, or vice versa—for example, "aku recommend" (I recommend) or "dream book aku semua" (all my dream books), where the English component occupies a position within an Indonesian construction. The category of "alternation" is applied when a comment exhibits a shift between relatively distinct English and Indonesian segments; for instance, a comment might begin with an English statement or expression, followed by an explanation in Indonesian. The determination of a dominant pattern is made at the level of the individual comment; thus, while a comment may contain multiple linguistic phenomena, only one primary pattern is assigned for the purpose of final tabulation.

The category of "congruent lexicalization" is considered when the data show that elements from two languages share or occupy the same grammatical structure in a more comprehensive manner—going beyond the simple insertion of a single element or the alternation between two segments. This category was not found in the final corpus. This does not imply that hybrid structures cannot occur in Indonesian–English comments; rather, it indicates that the analyzed data did not meet the operational criteria established for this study. Because Muysken's typology was originally formulated based on the study of code-mixing in spoken bilingual discourse, its application to written comments serves as an analytical adaptation, with careful attention paid to the boundaries of written phrases and clauses.

In identifying English elements, this study employs an operational criterion: any element that retains its English form while appearing within a construction that is predominantly in Indonesian is classified as code-mixing—even if some of these terms have become commonplace in digital communication. Words such as "preloved", "insight", "habit", "recommend", "relate", "down", and "overthinking" represent borderline cases that require consistent treatment. This criterion does not imply that these words have not been assimilated into the Indonesian language; rather, it clarifies the analytical decisions made within the study's corpus. Hybrid forms—such as "ternoticee", "melooting", and "kehandle"—are categorized as combinations of English elements and Indonesian morphology. This classification requires careful interpretation, as the boundaries between code-mixing, lexical borrowing, and digital slang are not always clearly defined.

Cross-linguistic analysis of linguistic elements requires careful attention to the form of a word and its function within the structural framework of the language in which it is used. Sitompul and Lingga (2024) compared the use of prefixes in Toba Batak and English using a qualitative descriptive approach, demonstrating differences in form and meaning within the word-formation systems of both languages. Although their study does not specifically address Indonesian–English code-mixing, it serves as a pertinent reminder that linguistic elements from different languages must be analyzed based on their respective forms and characteristics. This principle underscores the need for consistent operational criteria when identifying English words and hybrid forms within TikTok comments.

Communicative functions are analyzed using the framework of Appel and Muysken (1987), which outlines the functions of language within the context of code-switching and bilingual communication. A referential function is identified when a comment primarily conveys information about a book, genre, content, or an experience that can be described informatively. A directive function is identified when a comment requests an answer, confirmation, or recommendation, or encourages another user to take a specific action. An expressive function is employed when the core content conveys an evaluation, a preference, disappointment, an emotional response, or a personal reaction. A phatic function is identified when a significant part of a comment

consists of greetings, expressions of gratitude, or remarks intended to maintain social rapport. A metalinguistic function is used when a comment addresses language itself—such as through quotations or specific expressions—treating it as the object of focus. A poetic function is considered when the primary emphasis lies on the form of the message, wordplay, or the stylistic manipulation of language.

A single dominant function is assigned to each comment to ensure the total number of entries in the tabulation corresponds to the total number of data points. Additional functions are noted if a comment serves other clear purposes. The assignment of a function is based on the content of the comment and the context of the post, rather than solely on the presence of English words. For instance, the phrase “really recommended” within a comment that includes a personal reading experience is classified primarily as expressive, whereas a question asking for book recommendations is classified as directive. Some comments may simultaneously provide information and convey an attitude; in such cases, the rationale for selecting the dominant function is explained in the data matrix and the interpretation of the results.

The functional categories employed are interpretive in nature. The coding process is not intended to determine the psychological state or personal motives of the account holders, as such information cannot be definitively ascertained from public comments alone. In this study, the term “expressive” refers to the function of an utterance or text in conveying evaluations or feelings manifested within the message itself—it does not constitute a diagnosis or a conclusion regarding the user’s emotional state beyond the scope of the comment.

The analysis was conducted in six stages. First, the collected comments were transcribed without altering spelling, capitalization, or character repetitions—features characteristic of digital communication. Second, each comment was assigned an anonymous code in the format C1-K01, C1-K02, and so on, where 'C' denotes the content and 'K' indicates the sequence of the comment. Third, the comments were filtered by applying inclusion and exclusion criteria. Fourth, English elements were identified, and dominant patterns were determined using the categories of insertion, alternation, and congruent lexicalization. Fifth, the communicative functions were classified based on the context and the apparent purpose of each message. Sixth, the coding results were tabulated, compared across different posts, and interpreted by relating them to the theoretical framework and previous research.

To enhance traceability, the matrix records the comment ID, the original excerpt, the identified English language elements, structural patterns, dominant functions, the rationale for coding, and the selection status. Each primary comment is counted once within the distribution of patterns and dominant functions. While comments may contain multiple language elements or serve several communicative purposes, they are not counted repeatedly in the frequency tables. Consistency checks were conducted by re-examining the comment texts and the coding rationales against the established criteria. This study did not employ a second coder or calculate an inter-rater reliability coefficient; this is acknowledged as a limitation that could be addressed in future research.

Ethical considerations regarding data usage were upheld by excluding the account names of commenters from the analytical matrix and the results tables. The identities of content creators are mentioned only to the extent necessary to describe the source of the public content. Comment excerpts were selected judiciously to substantiate the coding and were not used to evaluate the individuals themselves. The access dates for each post were not recorded in the initial dataset and will need to be supplemented based on the original documentation. This limitation is significant, as the display,

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

sequence, and availability of comments on the platform may change over time.

HASIL DAN PEMBAHASAN

1. Corpus profile and filtering results

The research corpus is derived from three posts, each featuring a different context for book-related discussion. Content C1 focuses on Indonesian literary works read by the creator throughout 2025; the comments include requests for confirmation of titles, recommendations for other works, evaluations, and the identification of genres. Content C2 centers on Indonesian books that left a strong impression on the creator; the comments here consist of requests for recommendations, appreciation for the book selections, and expressions of interest in reading the featured titles. Content C3 discusses the book *Insecurity Is My Middle Name*; the comments tend to evaluate the book and connect its content to the readers' own experiences or feelings. This description is intended to provide context for interpretation, rather than to draw conclusions about the general characteristics of all users across these accounts.

Of the 45 comments examined, 34 (75.6%) were included in the primary dataset, while 11 (24.4%) were excluded based on operational criteria. The distribution varies by content: C1 yielded 12 primary comments, C2 yielded eight, and C3 yielded 14. This variation stems from the number of comments that met the criteria within the initial sample, rather than from a comprehensive comparison of all interactions across the three posts. Content C3 provided the largest share of the primary dataset, whereas C2 contained a higher number of English-language comments that merely mentioned Indonesian book titles as proper nouns.

Filtering is a crucial step because comments containing two languages do not necessarily exhibit a relevant code-switching structure. For instance, a comment written entirely in English that mentions the title of an Indonesian book does not automatically demonstrate the insertion of Indonesian into an English syntactic structure; the title may simply function as a proper noun. Conversely, a comment such as “aku pengen mulai belajar bangun habit baca buku, dari semua ini yg paling recommend yang mana ya?” features an English word within an Indonesian sentence structure, thus qualifying it for analysis according to the operational definition. These inclusion and exclusion decisions are intended to reduce the likelihood of counting titles of works as instances of code-mixing in the absence of sufficient structural evidence.

Table 1. Number of comments analyzed for each piece of content

Content code	Brief context	Comments transcribed	Enter analysis	Exclusion
C1	Indonesian Literature Reading List 2025	15	12	3
C2	Indonesian books that leave a deep impression	15	8	7
C3	The book “Insecurity Is My Middle Name”	15	14	1
Total	Three book-themed posts	45	34	11

2. Distribution of code-mixing patterns

The *insertion* pattern was found in 22 out of 34 primary comments (64.7%), while the *alternation* pattern was found in 12 comments (35.3%). No comments were classified as *congruent lexicalization*. Based on the coding

criteria used in this study, insertion emerged as the most frequent pattern, as the majority of the data featured English words, phrases, or expressions embedded within structures primarily built in Indonesian. The *alternation* pattern occurred when comments shifted between larger language segments—for instance, an English-language segment followed by an explanation or assessment in Indonesian.

This distribution varies across the different pieces of content. C1 contains nine instances of *insertion* and three of *alternation*; C2 contains six instances of *insertion* and two of *alternation*; and C3 contains seven instances of insertion and seven of alternation. Consequently, the overall dominance of insertion across the entire corpus is not uniform across all content: in the C3 sample, both patterns appear with equal frequency. This suggests that any interpretation of the total distribution must account for the specific context of each post—particularly given the differences in topics, content styles, and the nature of the comments collected.

Percentages are calculated based on 34 primary comments and rounded to one decimal place. Each comment counts as a single frequency unit, even if it contains multiple English elements or more than one type of mixed-language structure. Consequently, the totals for the categories in the table do not represent a count of every individual English word or phrase present in the corpus. While this approach ensures consistency across the comment units, it also means that the complexity within a single comment is reduced to a single dominant pattern.

2.1. Insertion: the inclusion of English elements within an Indonesian structure

An insertion pattern is observed in comments that maintain Indonesian as the primary framework, with English elements occupying specific positions within the sentences. An example can be found in C1-K03: “anyonee, Namaku Alam stand alone bukaan ya??” In this comment, the core question is formulated in Indonesian, while the term “anyonee” and the phrase “stand-alone” are derived from English. The term “stand-alone” also relates to book classification, as it is used to inquire whether a work is a self-contained story or part of a series. These English elements serve to convey specific meanings relevant to the context of a book discussion, even though the overall structure of the question remains dominated by Indonesian.

Comment C1-K06 demonstrates the use of several terms related to genre and publishing, such as “stand alone,” “historical,” “political,” “series,” “trilogy,” and “hisfic.” These elements appear within a comment that is otherwise in Indonesian, which includes book recommendations and descriptions of the stories' settings. This data suggests that code-switching can be linked to the use of concise, specific terminology common within a community of readers. However, this study cannot definitively determine whether these terms were chosen because no Indonesian equivalents exist, because the Indonesian alternatives are considered too lengthy, or simply because the English terms are more familiar to the commenters. Any justifiable interpretation is limited to the function of these terms within the text as presented.

A similar pattern appears in C2-K11, where the expressions “habit baca buku” and “recommend” are used in a query about books suitable for someone looking to start a reading habit. These English elements are embedded within an Indonesian sentence structure, clearly conveying the user’s intent: a request for reading recommendations. In C2-K13, the phrase “historical fiction” appears in a question regarding a new book titled “Katri”. Here, the phrase functions as a genre label, while the surrounding sentence structure—which poses the inquiry—remains in Indonesian. Both

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

examples demonstrate that such insertions are not limited to single words; English phrases can also function as elements occupying lexical positions within an Indonesian sentence.

In C3, the form of "insertion" is observed in C3-K03, which employs the word "relate" to connect the reading experience with personal circumstances; in C3-K10 and C3-K13, which use the phrase "really recommended" to evaluate or suggest a book; and in C3-K11, which incorporates the terms "relate" and "down" within a text that is otherwise predominantly in Indonesian. The forms "kehandle" (in C3-K11) and "melooting" (in C3-K07) also exhibit morphological hybridity: an English lexical base is combined with Indonesian word-formation elements or structural frameworks. For both of these hybrid forms, the term "insertion" is applied as the dominant pattern at the level of commentary—a classification that does not negate the mixed morphological nature inherent in the words themselves.

This category requires careful interpretation, as terms such as “preloved,” “insight,” “relate,” “habit,” and “recommend” have become commonplace in Indonesian digital communication. If a study were to apply stricter criteria regarding lexicalization or loanword integration, some of these elements might be excluded from the code-switching data. This research addresses the issue through a consistent operational decision applied to the entire corpus: forms that are written as English words and function as elements within an Indonesian sentence are recorded, whereas book titles or proper nouns—when they constitute the sole element from another language—do not meet the criteria. This operational choice must be stated explicitly, as it influences the total number of identified insertions.

Code-insertion can also serve to support an evaluative style. In C1-K14, the terms “bitter,” “dark,” “amaze,” and “niche” appear alongside Indonesian phrases, including the expression “and I love it.” While the comments as a whole cannot be categorized merely as lists of vocabulary, the use of English elements contributes to a series of brief assessments regarding the complexity and nuances of the stories. Similarly, the term “dream book” in C1-K13 is embedded within the Indonesian structure “ini dream book aku semua” signaling the reader’s desire for, or admiration of, the books being presented. Both examples demonstrate that code-insertion can function in tandem with an expressive purpose—specifically, in articulating personal tastes or interests.

2.2. Alternation: the switching between language segments

An alternation pattern was identified in 12 comments. One example is C1-K09: ““I also love Pulang by Leila. Jujur sama, semenjak baca buku ini. Jadi suka historical fiction”.This comment contains a statement in English, a segment in Indonesian, and then a genre term in English. Such a shift goes beyond the mere insertion of a single word into a sentence structure; therefore, the pattern is classified as an alternation, even though the same comment also includes an English phrase within the Indonesian portion.

C2-K09 demonstrates a clear shift from a question in Indonesian to a recommendation in English: “KAKK udh baca Kita Pergi Hari Ini blummm?? i think if u like Tanah Lada and Semua ikan di Langit u will like that tooo!” The first part of the comment seeks confirmation regarding the interlocutor’s reading, while the subsequent English clause offers a recommendation based on the interlocutor’s preferences or reading history. In this instance, English is not merely used for a genre term or an adjective; rather, it functions across a relatively complete sequence of clauses. This pattern illustrates a code-switching phenomenon that contributes to the organization of the message within the comment.

Another example can be found in C3-K01, , “i have this book, BENER BENER BAGUSSS”, as well as in C3-K05, which begins with an assessment in Indonesian before shifting to “i like the way u written that books, thank u.” Both comments employ segments of English and Indonesian to express an evaluation of the book. This language alternation does not necessarily signal a change in the communicative situation; in these instances, the content of the message remains centered on a single evaluation—namely, that the book is considered good. Consequently, such alternation within the corpus cannot be directly equated with a shift in the communicative context; rather, it is best described based on language segmentation at the level of the individual comment.

C1-K04, “should i read pulang tahun depan? walaupun masi cukup trauma sama mas biru”, is a more complex example. The initial part follows an English interrogative structure, yet it incorporates the book title “Pulang” and the phrase “tahun depan”, followed by an explanation in Indonesian. This data exhibits mixing at the phrasal level as well as shifts between segments. For the purpose of recapitulation, one dominant pattern is designated as "alternation," as shifts in language segments are evident throughout the commentary. However, this example also illustrates the limitations of a single-category typology: a brief comment may contain more than one mechanism of code-mixing.

The use of alternation must also be distinguished from comments that are entirely in English but happen to include the titles of Indonesian books. In comments excluded from the primary dataset, titles such as Pulang or Ayah function merely as the names of works within an English sentence and do not, in themselves, constitute a code-switching structure. In contrast, in instances such as C1-K09 or C2-K09, the shift between language segments is more clearly evident in the relationship between the Indonesian and English portions. This distinction makes structural criteria an essential factor in the selection process, although the classification of certain borderline cases remains open to debate.

2.3. Congruent lexicalization that was not found

This corpus does not contain any comments classified as congruent lexicalization. Within Muysken’s (2000) framework, this category refers to the lexical mixing of two languages within a shared or mutually compatible grammatical structure. The research data more frequently exhibit English elements that are identifiable as insertions or segments shifted from Indonesian into English. Hybrid forms such as “ternoticee,” “melooting,” or “kehandle” are not automatically classified under the category of congruent lexicalization; in this coding scheme, such forms are noted as hybrid lexical elements, with a comment indicating that the dominant pattern is insertion.

The absence of this category should be understood as a result of the coding process applied to 34 primary comments, rather than a conclusion that such a pattern never occurs in book-related commentary. The small size of the corpus, with individual comments serving as the units of analysis, limits the opportunity to identify more complex constructions. Furthermore, applying categories originally formulated in studies of bilingual code-switching to digital texts requires clear operational definitions. In future research, the boundaries between morphological insertions, borrowings, and congruent lexicalization could be analyzed using more detailed procedures or assessed by a second coder.

3. Compact matrix of master data and coding traces

Table 4 provides a summary of the 34 comments used in the analysis. Snippets are selected to show relevant passages and can use ellipsis to shorten long comments; the sign does

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

not indicate new data or words added by the researcher. The "pattern" column shows the dominant category at the comment level, while the function shows the categories used in the frequency table. The reasons concisely explain the basis of the classification without stating the personal motives of the commentator. These tables are intended to make coding decisions traceable, not as a replacement for the entire analysis log.

Table 2. Coding of comments on content C1: Indonesian literature reading list

ID	Excerpts of commentary and English elements	Dominant pattern	Dominant function	Coding basics
C1-K02	"nitip aku mau preloved ini, masih mulus baru 4 bulan beli" Element: preloved	Insertion	Directive	English terms appear within the Indonesian structure; their function relates to offers/transactions.
C1-K03	"anyonee, Namaku Alam stand alone bukaan ya??" Element: anyonee; stand alone	Insertion	Directive	The Indonesian query includes an English expression; it seeks confirmation regarding the status of the book.
C1-K04	"should i read pulang tahun depan? walaupun masi cukup trauma sama mas biru" Element: should i read; "pulang tahun depan" within the framework of the English question	Alternation	Directive	The English inquiry segment and the Indonesian explanation appear within a single comment; the primary function is to solicit advice.
...	...	...	...	...
C1-K15	"so far AKU SUKA BGT ENTROK" Element: so far	Insertion	Expressive	"So far" introduces an evaluative clause in Indonesian.

Table 3. Coding of comments on content C2: Indonesian books that left a deep impression

ID	Excerpts of commentary and English elements	Dominant pattern	Dominant function	Coding basics
C2-K04	"omaygat ayah ternoticee ... pemilihan kata ... bagusss" Element: omaygat; ternoticee (a hybrid form of notice + ter-)	Insertion	Expressive	Exclamations and English-based hybrid forms appear in Indonesian commentary.
C2-K06	"di tanah lada ... ga mau dibaca lagi krn TRAUMATIC" Element: traumatic	Insertion	Expressive	"Traumatic" appears in the Indonesian structure to convey the emotional impact of the book.
C2-K09	"... udh baca Kita Pergi Hari Ini ... i think if u like ... u will like that tooo!" Element: i think if u like; u will like that too	Alternation	Directive	An Indonesian question followed by an English recommendation clause.
...	...	...	...	...

C2-K14	"ayah one of the best sih" Element: one of the best	Insertion	Expressive	English evaluative phrases appear within Indonesian constructions and express appreciation.

Table 4. Coding of comments on content C3: Insecurity Is My Middle Name

ID	Excerpts of commentary and English elements	Dominant pattern	Dominant function	Coding basics
C3-K01	"i have this book, BENER BENER BAGUSSS" Element: i have this book; bener bener bagus	Alternation	Expressive	The English statement is followed by an assessment in Indonesian.
C3-K02	"Bukunya super duper bagus banget ... overthinking, dan loneliness" Element: super duper; overthinking; loneliness	Insertion	Expressive	English elements are inserted into Indonesian sentences that evaluate a text.
C3-K03	"gue baca buku ini ... nusuk banget dan relate sama keadaan gue" Element: relate	Insertion	Expressive	In the context of Indonesian sentence structures, "relate" is used to express the connection between experiences.
...	...	...	...	...
C3-K15	"Ya betul saya mengapresiasi perjalanan hidupmu ... ngga semua wanita seperti u" Element: u (possible texting abbreviation for you)	Insertion	Expressive	The abbreviation "u" is interpreted as a texting form of "you"; the decision depends on operational criteria.

3. Distribution of communicative functions

The expressive function is the dominant function in 23 comments (67.6%). The directive function appears in seven comments (20.6%); the referential function in two comments (5.9%); the phatic function in one comment (2.9%); and the metalinguistic function in one comment (2.9%). The poetic function is not identified as the dominant function in the primary data. Percentages are rounded to one decimal place, which may cause the total to differ slightly from 100%. These results indicate that the comments across the three posts consist more of evaluations, expressions of appreciation, statements of likes or dislikes, and reactions to the reading experience, rather than the mere conveyance of neutral information.

Interpreting the meaning of linguistic expressions also requires attention to the relationship between linguistic form and the meaning being conveyed. Nurlaily et al. (2026), in their study on conceptual metaphors in the novel "Your Blood, My Bones", identified the use of metaphors based on a cognitive semantic approach, demonstrating that metaphorical expressions can be linked to the understanding of concepts, including emotions and communication. As that study focused on literary texts rather than social media commentary, its findings cannot be directly applied to explain the function of code-mixing in

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

the present research. Nevertheless, the study offers a supporting perspective, suggesting that the interpretation of linguistic expressions must take into account the meaning constructed within the context of their use.

The expressive function is evident in comments that describe a book as "buku "bagus", "relate", "bitter", "dark", or "really recommended." English expressions frequently appear as part of concise, emphatic assessments. In C1-K14, for instance, terms like "bitter", "dark", "amaze", dan "I love it" accompany remarks about a story perceived as complex. In C3-K10, the phrase "really recommended" is used following an Indonesian-language comment that evaluates the book very positively. The English elements in these examples should not be viewed as a complete replacement for Indonesian; rather, their function emerges from their relationship with the surrounding Indonesian-language assessments.

The expressive function is also linked to personal experience. Comment C3-K03 notes that the book's content feels "relatable" to the commenter's own situation, while C3-K11 connects the experience of reading the book to a time when they were feeling "down," noting that it provided them with support. In both instances, the English language elements serve to frame the evaluation of the experiences being recounted. This interpretation is based on the form and content of the sentences, rather than an assessment of the users' emotional states beyond what is expressed in their published comments. With this limitation in mind, these comments can be viewed as examples of how language functions as a means to articulate a connection between the act of reading and personal evaluation.

The directive function is found when comments ask for confirmation, ask for recommendations, or invite the interlocutor to provide a response. C1-K03 uses the question form "bukaan ya??" after mentioning the term "stand alone"; C1-K04 asked whether it was necessary to read Pulang the following year; while C2-K11 asked which books are most recommended for people who want to build a reading habit. This directive function arises because comments are directed at creators or other users and explicitly open up the possibility of answers. Code mixing becomes part of the question or request, but the directive function is not determined just because English elements appear in it.

The referential function is found in two comments that primarily convey information. C1-K06 contains recommendations for various works, genres, and series statuses, while C1-K07 explains how a book can provide insight into grief. In these comments, English elements serve to identify specific concepts or terms used within the information being presented. The referential category must be distinguished from the directive category, as longer comments may simultaneously offer recommendations and provide information. For the purpose of frequency distribution, the function that is most prominent—or that serves as the primary objective of the message—is selected as the dominant category.

The phatic function is assigned to C2-K10; it begins with "thank you" followed by an explanation in Indonesian and the phrase "I will try for now." The expression of gratitude serves to acknowledge the preceding recommendation while maintaining the smooth flow of interaction between the user and the creator. Although the comment also contains a statement of intent to try the book, the phatic function is selected as the primary function for the purpose of tabulation, as the opening expression—which sustains the interactive relationship—constitutes a significant element at the start of the response. This classification is the result of a qualitative interpretation and can be verified by

another reviewer using the same set of guidelines.

The metalinguistic function is identified in C3-K12. The comment explicitly refers to the "quotes di atasnya" and cites the English expressions: "That you matter, that you're enough, that you complete". The focus is not merely on evaluating the book, but rather on the quotations contained within its content. Since the English segment consists of a direct quotation, this instance must be distinguished from the more productive code-switching found in a user's spontaneous speech. Nevertheless, the comment is included in the primary corpus according to the research criteria, as its structure combines Indonesian text with English expressions and explicitly centers the English text as the primary object of attention.

None of the 34 entries in the dataset can be classified as having a predominantly poetic function. This is to be expected in a small corpus focused on reader comments, which largely consist of recommendations, evaluations, and direct responses to books. While some comments do employ features such as capitalization, repeated letters, or emojis, these elements are not automatically categorized as serving a poetic function. The "poetic" category is applied only when the manipulation of the message's form itself stands out as the primary objective; the evidence within this corpus is insufficient to identify any single comment as a clear example of this function.

4. Relationship to previous research

The prevalence of the insertion pattern in this corpus aligns, in a general sense, with research indicating that TikTok comments can incorporate words, phrases, and clauses from other languages. Mustikarani and Pratiwi (2025), for instance, identified various forms of insertion within comments on the TikTok accounts they examined. Similarly, Sari et al. (2026) observed code-switching—specifically involving words and phrases—in comments within K-pop fan communities. While these parallels support the validity of using comments as a unit of linguistic analysis, they do not imply that the distribution of patterns in the present study is identical to that of previous research; factors such as the source of the content, data collection techniques, selection criteria, and the definitions of specific categories may vary.

The tendency toward an expressive function in this dataset can also be compared to the work of Tarigan (2026), who examined 25 mixed-language social media comments and identified several communicative functions, including expressive, supportive, persuasive, and solidarity-oriented functions. Sari et al. (2026) framed the use of code-mixing in K-pop fan comments as a means of expression, appreciation, identity, and community solidarity. Both studies demonstrate that comments can serve social functions beyond the mere conveyance of literal information. In the present study, the functional categories employed differ—being more closely aligned with the framework of Appel and Muysken—which necessitates a conceptual comparison rather than a direct, one-to-one mapping of every functional label.

The study of linguistic expressions also highlights the importance of considering the relationship between the form of an utterance and the meaning intended to be conveyed. Fardiyyah et al. (2025), in their research on methods for translating figurative language in the album "The Tortured Poets Department: The Anthology", identify various forms of figurative language and discuss how the meaning and structure of these expressions are rendered into Indonesian. Although their research focuses on the translation of song lyrics rather than bilingual interaction on social media, it serves as a relevant point of comparison for discussing the interpretation of expressions. In the present study, a similar principle of caution is applied by interpreting the function of comments based on

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

their content and context, rather than relying solely on the presence of English words or phrases.

Research by Tarihoran et al. (2022) and Rizky et al. (2024) provides a broader context regarding the use of—and attitudes toward—Indonesian–English code-switching on social media. Both studies utilize different data or pursue different objectives than the present article: Tarihoran et al. employed questionnaires and interviews with Generation Z, while Rizky et al. measured the attitudes of Twitter users. The current study does not collect data on user perceptions or motivations; therefore, it cannot replicate findings concerning positive attitudes or the reasons behind the choice to use English. Instead, its results complement this existing context by illustrating the specific forms present in actual comments, as well as the communicative functions that can be discerned from the content of the messages.

Astuti et al. (2023) highlight the significance of mixed-language data for the development of computational sentiment analysis on Twitter. Although their methodological orientation differs, their study underscores that Indonesian–English code-switched data possesses a structure that must be carefully considered during the annotation and processing stages. This type of research also relies on annotation decisions—such as whether a given text is classified as an "insertion," an "alternation," or falls outside the primary dataset—as well as the identification of the most dominant communicative function. Consequently, the transparent documentation of coding rules and illustrative examples is an essential component of reporting qualitative research.

Regarding book-related content, Christie et al. (2024) highlight themes of community and reader identity within #BookTokIndonesia videos, while Toisuta (2024) examines the characteristics of "netspeak" on BookTok and Bookstagram. The present study focuses on a smaller unit of analysis—specifically, the comments on three book-themed posts—rather than analyzing entire #BookTokIndonesia videos. Consequently, its contribution is more circumscribed, centering on a description of the structures of language mixing and the functions of the comments analyzed. The findings are not intended to suggest that the reader community employs a uniform linguistic pattern; rather, the variations observed across C1, C2, and C3 demonstrate that the context of each post must be taken into account when interpreting the data.

SIMPULAN

Based on 34 comments that met the criteria from three book-themed TikTok posts, two patterns of code-mixing were identified as dominant: insertion, appearing in 22 comments (64.7%), and alternation, appearing in 12 comments (35.3%). The congruent lexicalization pattern was not found in the analyzed corpus. Insertion was more prevalent in the combined dataset, although the composition varied by content: C1 and C2 contained a higher frequency of insertion, while C3 showed an equal number of instances of insertion and alternation. These results indicate that English elements within the comments do not always function as isolated, standalone words; some comments exhibit shifts involving broader segments of language.

Regarding communicative function, the expressive category was the most dominant, appearing in 23 comments (67.6%), followed by the directive function in seven comments (20.6%), the referential function in two comments (5.9%), and the phatic and metalinguistic functions in one comment each (2.9%). No comments were classified as having a predominantly poetic function. The dominance of the expressive function suggests that the comments in this corpus are largely used to evaluate books, express personal interest, and

connect the reading material to the readers' own experiences. The directive function manifests through questions, requests for confirmation, and reading recommendations. These findings are limited to the 34 selected comments and, therefore, cannot be generalized to the entire BookTok community or all TikTok users.

Future research is advised to expand the number of posts and the variety of source accounts, as well as to establish more detailed procedures for selecting comments. Researchers could specify the data collection period, the sequence of comments observed, the number of initial comments per post, and the date of access, thereby making the documentation process easier to replicate. Expanding the range of sources would also allow for comparisons between book review content, reading recommendations, responses to excerpts, and conversations within reader communities.

DAFTAR PUSTAKA

Amiroh, P. D., Ulfiyani, S., Mukhlis, & Utami, H. R. (2025). Code-switching and code-mixing in the language use of influencer @Linglinglieanti on TikTok. Konfiks: Jurnal Bahasa, Sastra, dan Pengajaran, 12(1), 199–211. https://doi.org/10.26618/rhp8xg03

Appel, R., & Muysken, P. (1987). Language contact and bilingualism. Edward Arnold.

Astuti, L. W., Sari, Y., & Suprapto. (2023). Code-mixed sentiment analysis using transformer for Twitter social media data. International Journal of Advanced Computer Science and Applications, 14(10), 498–504. https://doi.org/10.14569/IJACSA.2023.0141053

Christie, E. A., Wibowo, K. A., & Fuady, I. (2024). #BookTokIndonesia: Analisis konten komunitas pembaca buku pada media sosial TikTok. Shaut Al-Maktabah: Jurnal Perpustakaan, Arsip dan Dokumentasi, 16(1), 109–122. https://doi.org/10.37108/shaut.v16i1.765

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. Journal of Advanced Nursing, 62(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Fardiyyah, I., Supardi, M., Akmal, H., & Sa'diyah, S. (2025). The translation method of figurative language in The Tortured Poets Department: The Anthology album by Taylor Swift. Journal of Translation, Linguistics and Literature, 2(1), 12–27. https://doi.org/10.26714/translite.v2i1.742

https://repository.syekhnurjati.ac.id/14699/

Mosende, K. D., Ada, S. D., & Redita, A. (2024). TikTok's EduTok: A learning tool for grammar. Journal of Translation, Linguistics and Literature, 1(1), 14–29. https://doi.org/10.26714/translite.v1i1.324

Mustikarani, A., & Pratiwi, D. R. (2025). Penggunaan campur kode warganet pada akun TikTok @sabrinachairunnisa_: Kajian kualitatif deskriptif. Basastra: Jurnal Bahasa, Sastra, dan Pengajarannya, 13(1). https://doi.org/10.20961/basastra.v13i2.97728

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

Nasir, A. (2025). Sosiolinguistik. CV Cemerlang Publishing. https://www.cvcemerlangpublishing.com/sosiolinguistik Nurpadillah, V. (2023). Buku ajar sosiolinguistik: Pemilihan kode tutur. Zenius Publisher.

E-ISSN : 3064-2965
Volume 3 No. 1, Mei 2026

Nurlaily, S., Supardi, M., Saefudin, & Farkhan, M. (2026). Conceptual metaphors in Kelly Andrew's Your Blood, My Bones (2024). Journal of Translation, Linguistics and Literature, 3(1), 25–38. https://doi.org/10.26714/translite.v3i1.799

Rizky, M., Purwanti, I. T., & Aruan, R. (2024). Indonesian Twitter users language attitude towards English-Indonesian code-mixing. IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, 12(1), 50–67. https://doi.org/10.24256/ideas.v12i1.2340

Sari, N. P. K., Sukarno, & Wicaksana, M. F. (2026). Campur kode dalam kolom komentar penggemar K-pop pada aplikasi TikTok. Jurnal Genre (Bahasa, Sastra, dan Pembelajarannya), 8(2), 401–413. https://doi.org/10.26555/jg.v8i2.16268

Siagian, C. B., & Damanik, B. A. R. (2024). The language production of monolingual children and multilingual children: A comparative study of psycholinguistics. Journal of Translation, Linguistics and Literature, 1(2), 80–87. https://doi.org/10.26714/translite.v1i2.402

Sitompul, H. T., & Lingga, R. K. (2024). A contrastive analysis on prefix usage in Batak Toba and English: Morphological perspectives. Journal of Translation, Linguistics and Literature, 1(2), 88–96. https://doi.org/10.26714/translite.v1i2.398

Social Research, 9(3), 3670–3679. https://doi.org/10.54314/jssr.v9i3.6225

Susylowati, E., Zakiyah, F., Sandy, D. K., & Cicilia, V. D. (2024). Sosiolinguistik: Teori dan aplikasi. Penerbit Underline. Tarigan, K. E. (2026). Indoglish in selected Indonesian social media comments: A qualitative analysis. Journal of Science and

Tarihoran, N., Fachriyah, E., Tressyalina, & Sumirat, I. R. (2022). The impact of social media on the use of code mixing by Generation Z. International Journal of Interactive Mobile Technologies, 16(7), 54–69. https://doi.org/10.3991/ijim.v16i07.27659

Toisuta, E. G. (2024). Language evolution in bookish social media: A netspeak analysis of BookTok and Bookstagram. Huele: Journal of Applied Linguistics, Literature and Culture, 4(2), 55–70. https://doi.org/10.30598/huele.v4.i2.p55-70

Tricahyo, A. (2019). Sosiolinguistik: Kajian budaya dalam analisis bahasa. CV Nata Karya. https://repository.iainponorogo.ac.id/1538/

Yanti, E., Poetri, M. S., Huan, E., Manafe, N. R., Sako, Y., & Putra, E. D. (2024). Pengantar sosiolinguistik. Aikomedia Press. https://aikomedia.id/pengantar-sosiolinguistik