

An Analysis of Code Mixing in Windah Basudara's YouTube Livestream: Types, Functions, and Sociolinguistic Implications

Ayu Diyasefia^{1*}

¹ Universitas Islam Riau, Indonesia

✉ ayudiyasefia@student.uir.ac.id*

Article information:

Received : nov 2025

Revised : 26 jan 2026

Accepted : 30 march
2026

Abstract

Code mixing, the integration of linguistic elements from two or more languages within a single communicative event, has become a defining feature of digital discourse in multilingual communities. This study investigates the types, frequency, and communicative functions of code mixing in the YouTube livestreams of Windah Basudara, one of Indonesia's most prominent gaming content creators with over 16 million subscribers. Employing a qualitative content analysis approach grounded in Muysken's (2000) tripartite typology insertion, alternation, and congruent lexicalization the study analysed three purposively selected livestream videos, each exceeding one million views. A total of 155 code-mixing instances were identified and systematically coded. Findings reveal that insertion was the most dominant type ($n = 75$; 48.38%), followed by alternation ($n = 61$; 39.35%) and congruent lexicalization ($n = 19$; 12.25%). These patterns reflect deliberate sociolinguistic strategies serving functions of identity construction, emotional emphasis, audience engagement, and digital persona building. The study contributes to the growing field of digital sociolinguistics and offers implications for bilingual education, media literacy, and the study of language practices in online entertainment environments.

Keywords: code mixing; Muysken's typology; digital discourse; YouTube livestream; sociolinguistics; Indonesian English bilingualism

INTRODUCTION

Language serves as a fundamental tool in shaping communication, particularly in the digital era where content creation and online interaction have become integral to daily life. The widespread use of social media platforms and video-sharing services, such as YouTube, has significantly influenced linguistic practices, leading to the emergence of new language phenomena. One such phenomenon is code mixing, which refers to the integration of two or more languages within a single utterance or discourse (Muysken, 2000). This practice is commonly observed among bilingual or multilingual speakers and serves various communicative functions, including identity construction, emphasis, and audience engagement.

Recent studies have highlighted the increasing prevalence of code mixing in digital communication. Dewirahmadanirwati and Aditiawarman (2023) argue that code mixing is frequently employed by Indonesian youth on social media as a strategy to express identity and build social rapport. Similarly, Susanti et al. (2024) emphasize the role of code mixing in digital mass media, noting its implications for language learning, public communication, and linguistic adaptability in online contexts. In the broader global context, Myers-Scotton's (1993) Markedness Model suggests that speakers strategically select language codes to signal social meanings—a process that is especially visible in the performative, audience-directed communication characteristic of live streaming.

In the Indonesian digital landscape, one notable figure who exemplifies this phenomenon is Windah Basudara, a prominent YouTuber and gaming content creator with over 16 million subscribers as of 2025. His livestreams are characterised by spontaneous and humorous commentary, often delivered through a dynamic alternation between Bahasa Indonesia and English. This bilingual communication style not only enhances the entertainment value of his content but also fosters a closer parasocial connection with his predominantly adolescent, digitally literate audience. The selection of Windah Basudara as the subject of this study is theoretically motivated: his livestream format—real-time, unscripted, and audience-interactive—constitutes an authentic and ecologically valid site for examining naturally occurring code mixing in digital discourse, free from the editing that characterises pre-recorded content.

Despite the growing body of research on code mixing in Indonesian digital media (Afrayanti et al., 2021; Najuar et al., 2025; Zebua et al., 2025), a specific gap exists in the

literature regarding the systematic linguistic analysis of code mixing in live streaming contexts. Most existing studies have focused on social media text (e.g., Instagram, Twitter) or pre-recorded YouTube videos, leaving the spontaneous, interactive, and performative language of livestreams comparatively underexplored. Furthermore, while the frequency and types of code mixing in Indonesian digital communication have been documented, the precise communicative functions of each code-mixing type within the context of gaming livestreams have not been comprehensively theorised.

This study therefore addresses the following research questions, framed within Muysken's (2000) typological framework:

RQ1: What types of code mixing as classified by Muysken's (2000) tripartite typology (insertion, alternation, congruent lexicalization) occur in Windah Basudara's YouTube livestreams, and with what frequency?

RQ2: What communicative functions does each type of code mixing serve in the context of gaming livestream discourse?

RQ3: What sociolinguistic implications do these code-mixing patterns carry for the study of bilingual identity construction in Indonesian digital media?

By addressing these questions, this study aims to make a dual contribution: empirically, by providing a systematic frequency analysis of code mixing in a high-visibility Indonesian livestream context; and theoretically, by extending the application of Muysken's typology to the emergent genre of gaming livestream discourse. The findings are expected to have practical relevance for researchers in digital sociolinguistics, educators engaged in bilingual or EFL instruction, and media scholars studying the intersection of language, identity, and digital entertainment.

METHODS

Research Design

This study employs a qualitative content analysis design to systematically examine and interpret patterns of code mixing in Windah Basudara's YouTube livestreams. Qualitative content analysis was selected as the appropriate methodology because it enables researchers to identify, categorise, and interpret linguistic phenomena within their communicative context, rather than merely quantifying surface-level occurrences (Krippendorff, 2018). As Aspers and Corte (2021) affirm, qualitative research is particularly suited to exploratory, contextually

situated investigations of social and communicative behaviour conditions that characterise the real-time, multimodal environment of live streaming.

The theoretical framework guiding the analysis is Muysken's (2000) tripartite typology of code mixing, which distinguishes three structural types: insertion (embedding of lexical items or constituents from Language B into a matrix frame of Language A), alternation (switching between languages at clause or sentence boundaries), and congruent lexicalization (use of a shared syntactic structure with lexical items from both languages). This typology was selected over alternative frameworks (e.g., Poplack's 1980 code-switching taxonomy) because of its demonstrated applicability to Indonesian-English bilingual discourse and its precision in distinguishing intra-sentential from inter-sentential mixing.

Data Sources and Sampling

Data were collected from three of Windah Basudara's YouTube livestream recordings, selected through purposive sampling based on the following pre-defined criteria: (a) view count exceeding one million, to ensure high audience engagement and representativeness of his mainstream communicative style; (b) duration of at least 60 minutes, to yield sufficient instances of code mixing for systematic analysis; (c) content type classified as gaming livestream specifically involving interactive gameplay commentary to ensure consistency of genre and communicative context across data sources; and (d) recorded between January 2024 and January 2025, to capture his current linguistic practices. Videos featuring predominantly guest commentary or edited highlight compilations were excluded to preserve the authenticity and spontaneity of the data.

Data Collection and Coding Procedure

Each livestream recording was transcribed verbatim by the researcher, with attention to both the Indonesian and English segments of each utterance. Transcription conventions followed Jefferson's (2004) simplified notation system. Code-mixing instances were then identified through a two-stage procedure: (1) segmentation, in which all utterances containing elements from both Bahasa Indonesia and English were isolated; and (2) classification, in which each instance was categorised according to Muysken's (2000) typology.

Coding was conducted independently by the researcher and a second coder trained in sociolinguistics. To assess coding reliability, Cohen's kappa was calculated following the initial coding round ($\kappa = .81$), indicating strong intercoder agreement (Landis & Koch, 1977).

Discrepancies were resolved through discussion and reference to Muysken's (2000) original definitional criteria. No dedicated software was required for segmentation, as the analysis was performed manually using a structured coding sheet; however, frequency tabulation and percentage calculations were performed using Microsoft Excel.

RESULTS AND DISCUSSION

Results

Based on the qualitative content analysis of three of Windah Basudara's most-viewed livestream recordings, a total of **155** code-mixing instances were identified across all three videos combined. These instances were classified into three categories following Muysken's (2000) typology: insertion, alternation, and congruent lexicalization. The frequency distribution is presented in Table 1.

Table 1. Frequency and Percentage of Code-Mixing Types in Windah Basudara's Livestreams

Types of Code Mixing	Frequency	Percentage
Insertion	75	48.38%
Alternation	61	39.35%
Congruent Lexicalization	19	12.25%
Total	155	100%

As Table 1 demonstrates, insertion constitutes the dominant type of code mixing (48.38%), followed by alternation (39.35%), with congruent lexicalization appearing least frequently (12.25%). Each type is discussed in detail below, with illustrative examples drawn from the transcribed data.

Insertion (n = 75; 48.38%)

Insertion the embedding of single words or short phrases from English into an otherwise Indonesian syntactic frame was the most prevalent code-mixing type observed. Representative instances include expressions such as "Let's go," "Nice," "Oh my God," and "streaming terbaik" (literally: the best streaming). These insertions predominantly involve high-frequency, globally circulating digital expressions that have become lexically integrated into Indonesian youth vernacular.

The dominance of insertion is theoretically significant. From a Muysken (2000) perspective, insertion presupposes a clear matrix language (here, Bahasa Indonesia) that governs the syntactic frame, with English elements functioning as embedded constituents.

This asymmetry reflects the sociolinguistic reality of Windah's audience: predominantly Indonesian-speaking adolescents with varying but generally intermediate English proficiency, for whom English insertions function as affective amplifiers and markers of digital cosmopolitanism rather than as primary communicative vehicles.

Functionally, insertion served three overlapping purposes in the data. First, emotional intensification: expressions such as "Oh my God" or "Let's go" are deployed at moments of high emotional arousal during gameplay, amplifying affective expression beyond what is conventionally achieved in Indonesian. Second, audience alignment: the use of globally familiar gaming expressions creates a sense of shared cultural membership between streamer and audience. Third, digital identity signalling: as Najuar et al. (2025) note, insertion of English gaming terminology by Indonesian YouTubers functions as a marker of expertise and cultural currency within digital entertainment communities.

Example (Insertion): "Udah gitu aja, terus kita push lane, let's go!" [That's all, then we push the lane, let's go!]

Alternation (n = 61; 39.35%)

Alternation code mixing involving a complete switch from one language to another at a clause or sentence boundary was the second most frequent type, occurring 61 times (39.35%). Examples include utterances such as "Guys, I think this is too much" and "What the hell just happened?" delivered following Indonesian-medium discourse.

Unlike insertion, alternation involves a more equal and symmetrical engagement with both linguistic systems, presupposing a higher level of English proficiency in both producer and intended audience. In the context of Windah's livestream, alternation predominantly occurred during moments of spontaneous, high-affect commentary such as responding to unexpected in-game events or engaging in direct audience address suggesting that the switch to English functions as a performative resource for dramatisation and expressive directness.

This interpretation is consistent with Zebua et al. (2025), who argue that alternation in Indonesian digital discourse serves dual communicative and performative functions: communicatively, it enables precise expression of concepts that carry different pragmatic weight in English than in Indonesian; performatively, it enacts a bilingual cosmopolitan persona that resonates with the aspirational identities of the streamer's Gen-Z audience. From a Myers-Scotton (1993) perspective, the shift to English at emotionally charged moments

constitutes a marked linguistic choice that signals authenticity, spontaneity, and affective investment.

Example (Alternation): "Bro, kita udah hampir menang nih. What the hell just happened? Kok bisa kalah?" [Bro, we were almost winning. What the hell just happened? How did we lose?]

Congruent Lexicalization (n = 19; 12.25%)

Congruent lexicalization the use of lexical items from two languages within a syntactically shared or parallel structure was the least frequent type, appearing in 19 instances (12.25%). Examples from the data include phrases such as "content creator," "option menu," and "new game," deployed within otherwise Indonesian utterances but without the clear embedding relationship that characterises insertion.

The relative infrequency of congruent lexicalization reflects the structural divergence between Bahasa Indonesia and English at the syntactic level: unlike cognate language pairs (e.g., Spanish-Portuguese) where congruent structures are pervasive, Indonesian-English bilingualism provides limited syntactic overlap, restricting the domains in which congruent lexicalization naturally emerges. When it does occur, it is most prevalent in the domain of technology and digital media terminology a finding consistent with Adnyasuari et al. (2024), who similarly report that congruent lexicalization in Indonesian digital discourse clusters around technical, domain-specific vocabulary where English has become the dominant nomenclatural register globally.

Functionally, congruent lexicalization in Windah's speech signals domain expertise and institutional legitimacy within the gaming and content creation industry. The use of terms like "content creator" and "option menu" without Indonesian equivalents reflects the process of lexical borrowing, where English technical vocabulary has become fully conventionalised within Indonesian gaming discourse, to the point where it is no longer perceived as code mixing by participants themselves a phenomenon Backus (2005) terms 'categorical borrowing'.

Example (Congruent Lexicalization): "Kita buka option menu dulu sebelum mulai new game." [Let's open the option menu first before starting a new game.]

Sociolinguistic Implications

Taken together, the findings demonstrate that code mixing in Windah Basudara's livestreams is not a symptom of linguistic confusion or inadequate proficiency but rather a sophisticated, intentional, and functionally differentiated communicative strategy. This interpretation aligns with Afryanti et al.'s (2021) characterisation of code mixing in Indonesian YouTube discourse as a form of linguistic performativity a practice through which identity, social relationships, and cultural positioning are constructed and negotiated in real time.

The patterns observed carry three broader sociolinguistic implications. First, they illustrate the ongoing transformation of Indonesian linguistic norms under conditions of globalisation and digital mediatisation, wherein English is no longer experienced exclusively as a foreign language but as a resource within a shared bilingual repertoire. Second, they highlight the role of high-visibility digital content creators as *de facto* agents of language change, whose linguistic practices are observed, imitated, and normalised by millions of adolescent viewers. Third, the data suggest that the genre of gaming livestreaming constitutes a linguistically productive environment for the emergence and stabilisation of new bilingual conventions—a finding with direct relevance for researchers tracking the evolution of Indonesian English contact varieties.

These findings also carry implications for bilingual education in Indonesia. The naturalistic code-mixing practices documented here may serve as authentic pedagogical materials for EFL classrooms, enabling teachers to leverage students' existing digital language experiences as entry points for formal instruction in English pragmatics, register, and genre.

CONCLUSIONS

This study investigated the types, frequency, and communicative functions of code mixing in Windah Basudara's YouTube livestreams, guided by Muysken's (2000) tripartite typology. Analysis of 155 code-mixing instances across three purposively selected livestream videos revealed a clear hierarchical pattern: insertion was the most dominant type (48.38%), followed by alternation (39.35%) and congruent lexicalization (12.25%). Each type was found to serve distinct yet overlapping communicative functions emotional intensification, audience alignment, performative identity construction, and domain-specific expertise signalling rather than representing arbitrary or unsystematic language behaviour.

The study concludes that code mixing in the context of Indonesian gaming livestreams constitutes a dynamic and purposeful sociolinguistic practice, reflecting the increasingly hybrid linguistic repertoires of Indonesian youth navigating between local and global communicative norms. Far from representing linguistic interference or deficiency, code mixing functions as a marker of communicative competence, cultural capital, and digital identity in contemporary Indonesian online entertainment.

Future research should extend this analysis to comparative multi-streamer studies to assess whether the patterns identified here are idiosyncratic to Windah Basudara's communicative style or representative of broader trends in Indonesian gaming livestream discourse. Longitudinal studies tracking shifts in code-mixing patterns as the Indonesian gaming community evolves would also enrich understanding of language change in digital media contexts. Additionally, audience reception studies examining how viewers perceive, interpret, and potentially adopt the code-mixing practices of their favourite streamers would yield valuable insights into the role of parasocial influence in language diffusion—an underexplored dimension of digital sociolinguistics with significant implications for language policy, media literacy education, and the study of digital-era language contact.

REFERENCES

- Adnyasuari, N. P. D., Pratiwi, P. A. A. S., & Putra, I. K. S. (2024). Analysis of intra-lexical code-mixing on podcast YouTube channel. *ELYSIAN JOURNAL: English Literature, Linguistics and Translation Studies*, 4(2), 186–195. <https://doi.org/10.36733/elysian.v4i2.8764>
- Afryanti, R., Daud, B., & Muthalib, K. A. (2021). A study of code-switching and code-mixing used on YouTube channel: A comparison of Indonesian YouTubers. *English Education Journal*, 12(3), 495–510. <https://doi.org/10.24815/eej.v12i3.19166>
- Aspers, P., & Corte, U. (2021). What is qualitative in research? *Qualitative Sociology*, 44(4), 599–608. <https://doi.org/10.1007/s11133-021-09497-w>
- Backus, A. (2005). Codeswitching and language change: One thing leads to another? *International Journal of Bilingualism*, 9(3–4), 307–340. <https://doi.org/10.1177/13670069050090030601>
- Dewirahmadanirwati, D., & Aditiawarman, M. (2023). Code mixing in Indonesian youth social media communication: Identity and social rapport. *Journal of Language and Communication*, 10(1), 45–58.
- EIKI Journal of Effective Teaching Methods. (2024). Influence of video games on children's English language learning, 2(1). <https://doi.org/10.59652/jetm.v2i1.165>
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–31). John Benjamins.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.

- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press. <https://doi.org/10.1017/s0022226703272297>
- Myers-Scotton, C. (1993). *Social motivations for codeswitching: Evidence from Africa*. Oxford University Press.
- Najuar, A. C., Rahmatunisa, W., & Wihadi, M. (2025). Analysis of code mixing on the YouTube channels of viral Z-generation Indonesian YouTubers. *Journal of Applied Linguistics*, 4(2), 177–187. <https://doi.org/10.52622/joal.v4i2.300>
- Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics*, 18(7–8), 581–618.
- Sirilakshmi, Y., Gogoi, B. P., Bhuyan, N., & Bunkar, R. C. (2024). Content analysis in qualitative research: Importance and application. ResearchGate. <https://www.researchgate.net/publication/385973745>
- Susanti, R., et al. (2024). Code mixing in digital mass media: Implications for language learning and public communication. *Indonesian Journal of Applied Linguistics*, 13(2), 310–322.
- Zebua, Y., Munthe, L., Manik, S., & Suprayetno, E. (2025). Code-mixing of Indonesian and English on Instagram social media. *Journal of Applied Linguistics*, 4(2), 292–301. <https://doi.org/10.52622/joal.v4i2.361>