

Language Using of Code-Mixing in Captions and Comments on Instagram's Content Creators

Silfi Aliya Sakina¹, Sigit Haryanto²

Corresponding Author: a320220091@student.ums.ac.id

Universitas Muhammadiyah Surakarta, Indonesia

DOI: [10.35974/acuity.v11i1.4249](https://doi.org/10.35974/acuity.v11i1.4249)

Abstract

Digital communication has increased the use of code-mixing among social media users, such as on Instagram. This research examines Indonesian-English code-mixing found in both captions and comments of three popular Indonesian bilingual content creators. Using a descriptive qualitative approach, 50 instances of code-mixing were identified and analyzed over a 4-month observation period. Each instance was thoroughly examined and categorized according to Muysken's (2000) typology, which includes insertion, alternation, and congruent lexicalization. In addition, the underlying motivations for code-mixing were analyzed using Holmes' (1992) motivational framework. Results show that insertion form (40%), followed by alternation (32%) and congruent lexicalization (28%). Motivationally, prestige and affective expression were the most salient factors, indicating that English serves both as a marker of status and as a means to connect emotionally with followers. Unlike previous studies that focused only on captions or product-promotion discourse, this research combines the classification of code-mixing types (Muysken, 2000) and motivational analyses (Holmes, 1992) across both captions and comments, providing a more comprehensive understanding of bilingual interaction in digital communication.

Keywords: *Bilingual, Code-mixing, Instagram, Social media, Sociolinguistic*

INTRODUCTION

Language functions as a primary means of human communication, whether spoken or written, and involves the structured and conventional use of words. Generally, language can be categorized into formal and informal types. Formal language is typically employed in official or professional contexts, whereas informal language is used in casual or everyday interactions (Sutrisno & Ariesta, 2019). In today's digital era, linguistic practices have evolved alongside technological advancement, particularly through social media. Platforms such as Instagram are among the most popular social media platforms. It enables users to combine visual and linguistic expression, allowing bilingual speakers to display identity, intimacy, and creativity through code-mixing, the blending of linguistic elements from two or more languages within a single discourse.

Code-mixing on Instagram is often observed in captions and comments, where users frequently integrate words, phrases, or entire clauses from bilingual languages, particularly English and their native tongues. In Indonesia, for instance, the blending of the Indonesian and English languages has emerged as a widespread practice, particularly among younger social

media users and influencers. This trend does not emphasize sentence structure but rather highlights variations in language usage and linguistic evolution within society (Andre, 2018). On social media platforms, it not only bridges linguistic gaps but also encourages cultural exchange by blending multiple languages, dialects, and cultural references in everyday digital communication.

According to Muysken (2000), code-mixing can be categorized into three main types: insertion, where elements of one language are inserted into another; alternation, which involves switching between languages within a conversation; and congruent lexicalization, where elements from different languages are mixed within the same grammatical structure. Although code-mixing draws from more than one language, it typically operates within a dominant linguistic framework. The mixed code thus maintains its own communicative function and autonomy, yet remains rooted in the speaker's language structure. According to Mesthrie et al. (2000), these features are not used arbitrarily; instead, they reflect the speaker's engagement with specific social norms, cultural values, and community practices.

Phenomena like bilingual and multilingual discourse, including code-switching and code-mixing, frequently take place in social communication (Alhakim & Novitasari, 2020). Rijal & Cholsy (2022) found that language variations appear in the form of words, phrases, and sentences, often including non-standard elements such as abbreviations, creative word choices, and the insertion of foreign terms. Based on the study conducted by Putri et al. (2024), the use of language in code-mixing on social media, particularly in Instagram captions, reveals the dominance of a specific type of code-mixing according to Suwito's (1983) theory of classification. In another study by Zebua et al. (2025), it was emphasized that code-mixing is a dynamic and adaptable linguistic practice shaped by digital culture and the communicative purposes of users.

According to Risliyanti et al. (2019), code-mixing usage on various social media platforms has notably increased. Additionally, Alhusna and Febriana (2022) found that code-mixing in captions often exhibits characteristics typical of women's language and is less commonly used by men. Another study highlights the significant role of social media in identity formation among Indonesian Generation Z, revealing that these platforms serve as vital cultural arenas for young people to actively shape their identities through diverse linguistic strategies. Similarly, Tarihoran et al. (2022) examined the role of social media in promoting code-mixing, particularly among the younger generation, to use both English and Indonesian on social media platforms. Anggarini et. al. (2022) identified three key findings related to the use of code-mixing by leading Indonesian selebgrams in promoting products on social media. The study highlights how code-mixing supports persuasive and dynamic communication, making it a valuable tool for selebgrams in digital marketing contexts.

Therefore, although many previous studies have analyzed code-mixing, there has been no specific research examining the language use of code-mixing by Instagram content creators according to Muysken's typology, unlike previous research, which was particularly focused on Suwito's theory, and other platforms, such as WhatsApp groups and Facebook statuses. Most of these previous studies concentrated on caption-based communication or product promotion contexts, while less attention has been given to how both captions and comment interactions reflect linguistic structure and social motivation simultaneously. To address this gap, this

research aims to explore the classification of code-mixing types Muysken's (2000) with motivational analysis Holmes' (1992) on captions and comments from Instagram content creators.' This research facilitates a comprehensive examination of both structural language patterns and underlying sociolinguistic factors.

METHODS

This research employed an exploratory descriptive qualitative design to analyse the patterns and motivations of Indonesian-English code-mixing among Instagram content creators. According to Bogdan and Biklen (1998), qualitative research is descriptive in nature, emphasising words, texts, or images over numerical figures. This method is especially effective in exploring how individuals interpret and respond to their experiences within specific social, cultural, or ideological contexts. Furthermore, Bogdan and Biklen (1998) state that research subjects are individuals selected as data sources based on the objectives and focus of the study.

This research aimed to identify the types of code-mixing used in social media content creators and explore the motivations, specifically in captions and comments on Instagram. Data were collected from the Instagram accounts of three bilingual Instagram content creators, referred to as Creator A, Creator B, and Creator C, to maintain anonymity. Selection criteria included: 1) each had at least 100,000 followers to ensure a wide and active audience research, 2) they posted in both Indonesian and English (visible from captions, or comments), 3) posted at least once a week during the four-month observation period (April-July 2025), 4) their bilingual self-presentation (use of English and Indonesian in posts) indicated consistent language alternation behaviour, aligning with Holmes (2020), on sociolinguistic markers of bilingual identity. These criteria ensured that the selected creators were representative of Indonesian bilingual social media influencers.

The researcher conducted systematic observation of captions and comments appearing on the creators' captions and comments. Screenshots were taken of relevant instances containing code-mixing. Loanwords and proper nouns (e.g., brand names, hashtags) were excluded to avoid lexical bias. A total of 50 valid instances of code-mixing were identified for analysis. Each instance was categorised based on Muysken's (2000) typology: 1) insertion, 2) alternation, and 3) congruent lexicalisation. Data frequencies were then presented as percentages to reveal dominant trends. This method aligns with the views of Sarwono (2022) and Satpathy et al. (2023), who emphasise that using simple percentages can enhance the clarity and interpretation of qualitative data. After classification, the data were organised into a detailed presentation to support the drawing of conclusions.

To explore the motivation behind code mixing, Holmes' (1992) framework was applied, which includes: 1) need feeling motive, 2) prestige motive (to express status or education), 3) affective motive (to express emotion or intimacy), 4) being more informative, 5) making jokes, and 6) self-emotion expression. To analyse motivation by observing the creator's background and the content discussed, and then classify them. The overall data collection process was carefully designed to ensure that all relevant information was gathered accurately and that findings were summarised in descriptive tables and charts to highlight dominant types and motivations. This research analysed only three creators over four months, limited to Indonesian-English pairs. Motivations inferred from text, not confirmed through interviews.

RESULTS

A total of 50 instances of Indonesian-English code-mixing were identified across captions and comments from the Creator A, Creator B, and Creator C consistently used code-mixing in their social media activity. Each instance was classified according to Muysken's typology into insertion, alternation, and congruent lexicalization.

Table 1. Types of Code-Mixing

No	Type of Code Mixing	Creator A	Creator B	Creator C	Frequencies	Percentage
1	Insertion	8	8	4	20	40.0%
2	Alternation	6	3	7	16	32.0%
3	Congruent Lexicalization	4	7	3	14	28.0%
Total		18	18	14	50	100%

Source: Data Adapted from Instagram Content Creators

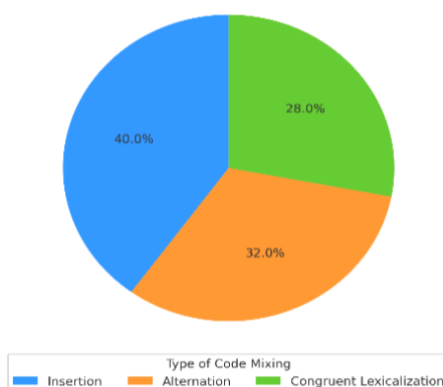


Figure 1: Pie Chart of Type of Code-Mixing

Source: Putri, 2024

The figure 1 presents the proportional distribution, showing insertion as the most frequent form of code-mixing among the three creators. To illustrate, the following excerpts show representative examples for each type: 1) insertion: “*sulit tapi bikin kamu grow*”; 2) alternation: “small things matter, *bikin simpel, bikin mudah dilakuin aja*”; 3) congruent lexicalization: *bangun tidur dan* feel good about yourself, *semua itu juga hak kamu.*” These examples show how English words or phrases are used in Indonesian sentences in different ways, with varying levels of grammatical blending. Based on quantitatively, insertion emerged as the dominant type (40%), followed by alternation (32%) and congruent lexicalization (28%). This distribution provides a descriptive overview of code-mixing patterns observed in the dataset not statistical data.

DISCUSSION

The data collected from 50 instances of code-mixing reveals that insertion was the most dominant type (40%), followed by alternation (32%) and congruent lexicalisation (28%). These results indicate that the three creators frequently inserted English words or short phrases into Indonesian sentences rather than switching entire clauses or blending grammar from both languages.

The predominance of insertion aligns with findings by Putri et al. (2024) and Risliyanti et al. (2019), who observed that Instagram users tend to use simple lexical insertions due to their practicality and recognizability. This pattern reflects the flexible linguistic behaviour of bilingual users in informal digital contexts. Insertion allows creator A to adopt English, such as “*sulit tapi bikin kamu GROW*,” another caption by Creator B “*, tuliss manifesting hal baru yang mau prenie lakuin itu apaaa???*” and “*kalo kamu wish-nya apa prenie?*” From these captions, the creators use words like ‘*grow, manifesting, wish*’ to add style, without disrupting Indonesian syntax. In contrast, alternation and congruent lexicalisation need more bilingual competence and are less common in brief social media captions.

Alternation refers to the combination of two distinct grammatical structures from different languages within a single sentence (Andre, 2018). In this dataset, it appeared mostly in extended captions that combine motivational or reflective statements, such as the captions of creator A, “loving the last sentence! *Bukan buat menghakimi, tapi buat bertumbuh.*” Another caption by creator B, “small things matter, *bikin simpel, bikin sederhana, bikin mudah dilakuin aja.*” This alternation usually helps move from one part of the caption to another, signalling shifts in tone or perspective. Congruent lexicalisation, while least frequent, refers to the blending of two languages within a single sentence, where both languages share a common grammatical structure (Andre, 2018). It shows deeper integration of the two languages, such as “*untuk setiap harinya bangun tidur dan feel good about yourself and your surroundings, semua itu juga hak kamu!*” from creator A. In other caption from creator B, “*belajar buat ‘no compare to others’ karena pada dasarnya tiap manusia punya lebih kurangnya masing-masing.*” This pattern suggests a creative linguistic identity consistent with the globalised nature of digital communication.

From the observation, according to Holmes’ (1992) motivational framework, two dominant motives emerged across the dataset: the prestige motive and the affective motive. Prestige motive appeared when creators used English to express educational professionalism or global orientation. For example, creator C frequently incorporated English expressions related to education and self-growth (“we launch opportunity,” “take this story as a reminder”). This shows they care about social status and having a worldly image. Using English in this way means it’s a symbol of modernity and upward mobility in Indonesian digital spaces. Affective motive was evident in emotionally charged captions that combined Indonesian empathy with English (e.g., *It’s okay to take a break* tapi jangan sampe satu momen sedih di hidup bikin kamu berhenti.”) Its English phrases reinforce emotional sincerity and intimacy with followers. The use of code-mixing also reflects a need-feeling motive. Especially, creator A often frames posts about mental health advice and motivation, which may indicate expertise, like “impostor syndrome,” “self-healing”, and “validation.” These lexical choices make the messages easier to understand and feel more personal to those who are familiar with digital audiences.

CONCLUSION

This research examined Indonesian-English code-mixing in the captions and comments of three Instagram content creators. Based on Muysken's (2000) typology, three types of code-mixing were found: insertion, alternation, and congruent lexicalization. Given the limited dataset, the findings are interpreted as exploratory rather than conclusive. The data identified 50 instances of code-mixing within the Instagram captions and comments. Of these, insertion was the most dominant form (40.0%), as it allows efficient integration of English words into Indonesian sentences. From Holmes' (1992) theory, the prestige and affective motives emerged as the key reasons for code-mixing, revealing that creators use English not only to convey information but also to express emotion and enhance their social image.

Overall, the findings highlight that code-mixing on Instagram functions as a sociolinguistic strategy that supports identity construction, emotional engagement, and communicative creativity in digital spaces. The practice illustrates how bilingual Indonesians navigate global and local identities simultaneously. The results support the notion of code-mixing as a marker of bilingual identity (Mesthrie et al., 2000; Tarihoran et al., 2022), where users negotiate between local and global cultural capital. Insertion serves as an accessible gateway for audiences, while alternation and congruent lexicalization demonstrate creative linguistic hybridity typical of Gen-Z digital communication.

This study analyzed only three bilingual creators within a limited four-month period, focusing solely on textual data without interview validation. Future research may expand the dataset, include audience perception, or compare code-mixing across different platforms to provide broader insights into bilingual digital discourse. Pedagogically, the findings contribute to sociolinguistic understanding of bilingual communication in Indonesia and suggest that educators may use authentic social media content to teach real-world code-mixing awareness in the English language classroom.

REFERENCES

- Alhusna, A., & Febriana, N. (2022). Code Mixing as Woman Identity: A Case on Instagram Captions of Influencer, *Jurnal Ilmiah Langue and Parole*, vol.6, no.1, p.1-6, December 2022. Available: doi.org/10.36057/jilp.v6i1.542
- Andre, R. (2018). Bahasa Indonesia-English Code Mixing in Writing Business Emails: Understanding the Communicative Purpose. *Lingua Cultura* 12(2):117–21. doi.org/10.21512/lc.v12i2.3733.
- Anggarini, K.T., Putra, I.N.A.J., & Suwastini, N.KA. (2022). An Analysis of Code-Mixing Used by The Indonesian Top Selebgram in Social-Media in Product Endorsement. *Jurnal Pendidikan Bahasa Inggris Indonesia* 10(1):54–65. [doi:10.23887/jpbi.v10i1.886](https://doi.org/10.23887/jpbi.v10i1.886).
- Bloomfield, L. (1984). *Language*. 2nd ed. New York: The University Chicago Press.
- Bogdan, R., & Biklen, K. S. (1998). *Qualitative Research for Education: An Introduction to Theory and Methods*. Allyn and Bacon.
- Coupland, N., & Jaworski, A. (1997). *Sociolinguistics: A Reader and Coursebook*. edited by N. Coupland and A. Jaworski. England: Palgrave Macmillan.
- Holmes, J. (1992). *An Introduction to Sociolinguistics*. London: Longman.
- Holmes, J. (2020). *An Introduction to Sociolinguistics*. 4th ed. edited by G. Leech and M. Short. New York: Routledge.

- Mesthrie, R., Swann, J., Deumert, A., & Leap, W. L. (2000). *Introducing Sociolinguistics*. Edinburg: Edinburg University Press.
- Muysken, P. (2000). *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- Putri, D.A.A.M., Padmadewi, N.N., Artini, L.P. (2024). Identifying Code-Mixing by Influencers in Captions on Social Media: Instagram. *Journal of Social Sciences and Humanities* 14(3). <https://doi.org/10.31940/soshum.v14i3.235-240>
- Rijal, A. S., Suhandano, & Cholsy, H., (2022). The Language Use on Instagram: Sociolinguistic Studies on the Speech Code of Netizens in the City of Makassar in Buying and Selling Food through Instagram. *AMCA Journal of Community Development* 2(2):34–41. [doi:10.51773/ajcd.v2i2.143](https://doi.org/10.51773/ajcd.v2i2.143).
- Risliyanti, I., Sitepu, S.S.W., & Tampubolon, M.A. (2019). Code-Mixing in Indonesian Selebgram's Caption in Instagram. *Lexeme: Journal of Linguistics and Applied Linguistics*. [doi:10.32493/ljal.v1i2.5126](https://doi.org/10.32493/ljal.v1i2.5126)
- Sarwono, J. (2022). *Quantitative, Qualitative and Mixed Method Research Methodology*. Seattle, WA: Amazon.com, Inc. ISBN 978-1076287885.
- Satpathy, B., Saha, S., Barman, B., & Roy, P., (2023). An Introduction to Research Methods and Designs in Social Sciences. *Advanced in Agricultural Research Methodology* pp. 7-20. India. ISBN 978-93-5780-711-1.
- Sutrisno, B., & Ariesta, Y. (2019). Beyond the Use of Code Mixing by Social Media Influencers in Instagram. *Advances in Language and Literary Studies* 10(6):143. [doi:10.7575/aiac.alls.v.10n.6p.143](https://doi.org/10.7575/aiac.alls.v.10n.6p.143).
- 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. [doi:10.3991/ijim.v16i07.27659](https://doi.org/10.3991/ijim.v16i07.27659).
- Wardhaugh, R. (2015). *An Introduction to Sociolinguistics*. UK: Blackwell Publishing.
- 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. [doi:10.52622/joal.v4i2.361](https://doi.org/10.52622/joal.v4i2.361).