

Code-Mixing in Agatha Chelsea's Utterances on Raditya Dika's YouTube Podcast

Campur Kode dalam Tuturan Agatha Chelsea di Podcast YouTube Raditya Dika

Sundari Daratista Nababan¹, Arsen Nahum Pasaribu², Rony Arahta Sembiring³

^{1,2,3}Universitas HKBP Nommensen Medan

sundari.nababan@student.uhn.ac.id , arsen.pasaribu@uhn.ac.id ,
ronyarahta.sembiring@uhn.ac.id

Article Information	ABSTRACT
<p>Description Submission: 18 March 2026 Revised: 22 April 2026 Accepted: 28 April 2026 Approved: 01 May 2026</p> <p>Keywords Code-Mixing Muysken's Typology Indonesian-English Bilingualism Podcast Conversational Context</p> <p>Kata Kunci Campur Kode Tipologi Muysken Bilingualisme Indonesia-Inggris Podcast Konteks Percakapan</p>	<p>This study aims to analyze the types and contextual patterns of code-mixing in the utterances of Agatha Chelsea on Raditya Dika's YouTube podcast using Muysken's (2000) typology. This study employs a descriptive qualitative method. The data consist of 173 utterances containing Indonesian-English code-mixing, collected through documentation and transcription, while data analysis is conducted using coding based on Muysken's structural classification and contextual observation. The results show that insertion is the dominant type (67.6%), followed by alternation (15.0%) and congruent lexicalization (17.4%). In terms of conversational contexts, insertion functions as a basic strategy across all settings. Alternation appears most frequently in interpersonal and relational contexts, where it marks shifts in perspective. Congruent lexicalization is prominent in academic and scientific discourse as well as in professional and creative practice, reflecting closer structural integration between the two languages. These findings indicate that code-mixing in Agatha Chelsea's speech is systematic and varies according to conversational context.</p> <p>Abstrak <i>Penelitian ini bertujuan untuk menganalisis tipe dan pola kontekstual campur kode dalam tuturan Agatha Chelsea pada podcast YouTube Raditya Dika menggunakan tipologi Muysken (2000). Penelitian ini menggunakan pendekatan deskriptif kualitatif. Data berupa 173 ujaran yang mengandung campur kode bahasa Indonesia-Inggris dikumpulkan melalui teknik dokumentasi dan transkripsi, sedangkan analisis data dilakukan dengan teknik pengkodean berdasarkan klasifikasi struktural Muysken dan observasi kontekstual. Hasil penelitian menunjukkan bahwa insertion merupakan jenis yang paling dominan (67,6%), diikuti oleh alternation (15,0%) dan congruent lexicalization (17,4%). Dalam konteks percakapan, insertion berfungsi sebagai strategi dasar di seluruh situasi. Alternation paling banyak ditemukan dalam konteks interpersonal dan relasional, berfungsi menandai pergeseran perspektif. Congruent lexicalization muncul secara menonjol dalam wacana akademis dan ilmiah serta praktik profesional dan kreatif, mencerminkan integrasi struktural yang lebih erat antara kedua bahasa. Temuan ini menunjukkan bahwa campur kode dalam tuturan Agatha Chelsea bersifat sistematis dan bervariasi sesuai konteks percakapan.</i></p>



Copyright © 2026 Sundari Daratista Nababan, Arsen Nahum Pasaribu, Rony Arahta Sembiring

1. Introduction

Language serves not only as a medium of communication but also as a reflection of social identity, cultural contact, and cognitive flexibility (Wardhaugh & Fuller. M, 2015). In bilingual or multilingual societies, the interaction between languages often gives rise to linguistic phenomena such as code-mixing, defined as the embedding of elements from one language into the grammatical structure of another within a single utterance (Muysken, 2000). Indonesia, with its more than 700 regional languages alongside Bahasa Indonesia as the national language, provides a fertile ground for such linguistic contact. In recent decades, the influence of English has grown significantly due to globalization, formal education, and the rapid expansion of digital media, resulting in widespread Indonesian–English bilingualism, particularly among urban, educated populations (Aditiawarman et al., 2025); (Hermawan et al., 2025). This bilingual environment has made code-mixing a common communicative strategy in both informal and semi-formal interactions.

The proliferation of digital media, especially podcasts and YouTube-based talk shows, has created new spaces for observing spontaneous, unscripted language use. Podcasts, characterized by extended conversational formats and minimal editing, allow speakers to produce natural speech that reflects authentic linguistic practices (Androutsopoulos, 2014). The Indonesian internet landscape further reinforces this trend; according to the Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), internet penetration reached 80.66% in 2025, with video-based platforms such as YouTube dominating long-form content consumption (Unit Survey APJII, 2025). Consequently, podcast discourse has become a significant site for sociolinguistic inquiry into contemporary bilingual practices.

Previous studies on code-mixing in Indonesia have examined various contexts, including classroom interactions (Sitinjak et al., 2020), social media comments (Safitri, 2021), and digital podcasts (Situmorang et al., 2023); (Saragih et al., 2025). Many of these studies employ Muysken's (2000) typology, which classifies code-mixing into three structural types: insertion, alternation, and congruent lexicalization. However, the majority focus on identifying and quantifying these types without systematically linking them to the conversational contexts in which they occur. In sociolinguistic terms, language variation is intrinsically tied to situational factors such as topic, participant roles, and communicative goals (Wardhaugh & Fuller. M, 2015). The absence of integrated analysis that considers both structural types and situational contexts represents a gap in the existing literature.

This study addresses that gap by focusing on the utterances of Agatha Chelsea, a young Indonesian public figure with a background in psychology and international exposure, in a podcast episode hosted by Raditya Dika titled "Cewek Cantik Ini Bisa Baca Orang?". The episode is selected because it contains a wide range of conversational topics—ranging from academic discussions of psychology and neuroscience to personal narratives, interpersonal relationships, professional experiences in acting, and casual interactions—thus offering a rich corpus for examining how code-mixing varies across contexts. Using Muysken's (2000)

typology as the analytical framework, this study aims (1) to identify the types of code-mixing present in Agatha Chelsea's utterances, and (2) to describe the patterns of code-mixing that emerge across different conversational contexts. Theoretically, this research contributes to the field of sociolinguistics by applying a structural typology to context-sensitive data, thereby enriching the understanding of code-mixing as a systematic and context-dependent phenomenon. Practically, it provides a model for analyzing bilingual speech in digital media and offers insights for educators, content creators, and researchers interested in contemporary Indonesian-English language contact.

2. Methods

This research employs a descriptive qualitative method to analyze code-mixing in the utterances of Agatha Chelsea on Raditya Dika's YouTube podcast. According to Creswell, (2009), qualitative research is an approach for exploring and understanding the meaning individuals or groups ascribe to social or human problems. Using this approach, the present study focuses on the in-depth analysis of naturally occurring spoken data in order to capture how code-mixing operates as a systematic and context-dependent linguistic phenomenon. This study aims to identify the types of code-mixing based on Muysken's (2000) typology and to examine how these types are distributed across different conversational contexts.

The data source of this research is a podcast episode uploaded on Raditya Dika's official YouTube channel entitled "Cewek Cantik Ini Bisa Baca Orang?". This episode was selected because it presents an extended and spontaneous interaction, allowing for the natural occurrence of Indonesian-English code-mixing. The data were collected using purposive sampling, focusing specifically on utterances produced by Agatha Chelsea that contain elements of code-mixing. The dataset consists of 173 utterances, each of which includes the insertion, alternation, or integration of English elements within Indonesian speech.

In this study, each utterance serves as the unit of analysis. An utterance is defined as a complete stretch of speech produced by the speaker, which may consist of a word, phrase, clause, or sentence containing code-mixing. Only utterances that exhibit Indonesian-English code-mixing were included in the analysis, while utterances without any mixing were excluded.

The data collection and analysis were conducted through several systematic steps:

a. Data Transcription and Selection

The podcast video was transcribed verbatim from the 51-minute recording. The transcription process was carried out through repeated viewing and careful listening to ensure accuracy. After the transcription was completed, all utterances containing Indonesian-English code-mixing were identified and selected. Utterances without any instances of code-mixing were excluded from the dataset.

b. Data Classification

The selected utterances were classified into three types of code-mixing based on Muysken's (2000) typology: insertion, alternation, and congruent lexicalization. In addition, the data were grouped into five conversational

contexts, namely: (1) academic and scientific discourse, (2) personal narrative and daily life, (3) interpersonal and relational discourse, (4) professional and creative practice, and (5) interaction and response. These categories were derived inductively based on recurring themes and communicative functions observed in the data.

c. Data Analysis and Interpretation

The classified data were analyzed to identify patterns in the distribution of code-mixing types across different contexts. The analysis focused on determining dominant types and explaining their communicative functions within each context. To ensure the trustworthiness of the findings, repeated examination of the data was conducted, including cross-checking the transcription and classification against the original podcast recording. This process aimed to ensure that the interpretations accurately reflected the actual conversational context.

3. Result and Discussions

This section presents the findings of the analysis of code-mixing in Agatha Chelsea’s utterances on Raditya Dika’s YouTube podcast. The results are organized based on the two research objectives: first, to identify the types of code-mixing using Muysken’s (2000) typology, and second, to describe the patterns of code-mixing across different conversational contexts. The analysis begins with a quantitative overview of the distribution of code-mixing types, followed by a contextual examination of how these types manifest in five distinct conversational situations. Each finding is supported by authentic examples drawn from the data and discussed in relation to existing theories and previous studies.

Types of Code-Mixing

From the 173 utterances analyzed, three types of code-mixing were identified based on Muysken’s (2000) classification: insertion, alternation, and congruent lexicalization. Table 1 presents the distribution of these types.

Table 1
Distribution of Code-Mixing Types

Code-Mixing Type	Frequency	Percentage
Insertion	117	67.6%
Alternation	26	15.0%
Congruent Lexicalization	30	17.4%
Total	173	100%

Insertion is the most dominant type (67.6%). This type is characterized by the embedding of English lexical items—such as words, phrases, or even affixed forms—into the grammatical structure of Indonesian, which serves as the matrix language. Examples include:

- 1) “Cuma *say hi* aja sih waktu itu.” – An English verb phrase is inserted into an Indonesian sentence.

- 2) "Kayak **personality-nya**." – An English noun combined with the Indonesian suffix *-nya*, showing morphological integration.
- 3) "Cuma bisa **di-replace**." – The English verb *replace* is combined with the Indonesian passive prefix *di-*, demonstrating productive morphological adaptation.

Alternation accounts for 15.0% of the data. This type involves switching between Indonesian and English at clause or phrase boundaries, with each language maintaining its own grammatical structure. Examples include:

- 1) "Dan ternyata aku belajar, **yes, and it happens** di setiap detik kehidupan kita." – A full English clause appears between Indonesian clauses.
- 2) "Jadi misalnya aku latihan nyanyi dari kecil, **and I do it all the time**, itu pasti otaknya jadi mengingat..." – Alternation marks a shift in perspective.
- 3) "Aku tetap mikir, **I have to be rational and thinking**." – An Indonesian clause is followed by an English clause that stands independently.

Congruent Lexicalization constitutes 17.4% of the data. This type reflects a high degree of structural integration, where lexical items from both languages are combined within a shared grammatical frame. Examples include:

- 1) "Jadi sebenarnya konsepnya adalah kalau misalnya neuron yang **fire together, they wire together**." – English technical expressions are seamlessly embedded into an Indonesian syntactic structure.
- 2) "**And it's easier** buat mereka **fire** lagi ketika kita melakukan kegiatan yang sama di masa depan." – English and Indonesian elements are interwoven within a single utterance.
- 3) "Makanya **one of the things that I like from acting** tuh itu sebenarnya ngebangun karakter itu." – A complex English clause is integrated with Indonesian particles *tuh* and *itu*.

Patterns of Code-Mixing across Conversational Contexts

Beyond structural classification, the analysis examined how code-mixing patterns vary across different conversational contexts. Based on thematic content and communicative functions, five contexts were identified. Table 2 presents the distribution of code-mixing types across these contexts.

Table 2
Distribution of Code-Mixing Types per Conversational Context

Conversational Context	Insertion	Alternation	Congruent Lexicalization	Total
Academic & Scientific	31	8	15	54
Personal Narrative & Daily Life	40	3	4	47
Interpersonal & Relational	24	12	5	41
Professional & Creative Practice	7	2	5	14
Interaction & Response	15	1	1	17
Total	117	26	30	173

Academic and Scientific Discourse

In this context, insertion is the most frequent type (31 occurrences), but congruent lexicalization appears with the highest proportion (15 out of 30 total instances). This indicates that when discussing complex topics such as psychology and neuroscience, the speaker integrates English technical terms and even entire clauses into Indonesian structures to maintain terminological accuracy and conceptual fluency. Examples:

- 1) "Kalau di psikologi kan ada yang **agreeableness**." – English term agreeableness inserted into Indonesian frame.
- 2) "Eh, sebenarnya kalau menurut neuroscience, kan addiction itu semuanya karena dopamin **reward pathway**, kan?" – Multiple English technical terms inserted.
- 3) "Jadi kita **unconsciously, the brain** itu selalu..." – English adverb and noun phrase inserted while maintaining Indonesian matrix.

Personal Narrative and Daily Life

Insertion overwhelmingly dominates (40 out of 47 utterances). The speaker maintains a stable Indonesian grammatical structure, inserting common English words and phrases that are widely used in informal Indonesian speech. Examples:

- 1) "Jadi **fun fact**, aku lagi masak sandwich waktu itu." – English fun fact and sandwich inserted.
- 2) "Nah, **the thing... fire alarm**-nya itu sensitif banget." – English phrase the thing and noun fire alarm with Indonesian suffix.
- 3) "Melbourne itu lebih **chill** malah menurut aku." – English adjective chill inserted into Indonesian comparative structure.

Interpersonal and Relational Discourse

Alternation reaches its highest frequency (12 out of 26 total alternation instances). The speaker shifts between Indonesian and English to differentiate between general statements and personal reflections. Examples:

- 1) "Karena perempuan itu **we really idealize romance**." – Switch to English after Indonesian clause to emphasize subjective opinion.
- 2) "Kalau misalnya enggak berkualitas, **it doesn't satisfied our needs for connection and relationship**..." – English clause inserted to explain the consequence.
- 3) "Jadi kita **tend to fall for a potential, instead of the actual person**." – English phrase inserted to describe a psychological tendency.

Professional and Creative Practice

Congruent lexicalization appears with notable frequency (5 out of 14 utterances). This context includes discussions about acting and character development, where the speaker blends both languages to express complex professional experiences. Examples:

- 1) "Makanya **one of the things that I like from acting** tuh itu sebenarnya ngebangun karakter itu." – English clause integrated with Indonesian particles.
- 2) "Karena cara aku untuk **solve problem itu, I have to understand things**." – Alternation combined with congruent lexicalization in professional problem-solving.

- 3) "Dan itu **step by step, nicotine intake-nya** itu **di-reduce** pelan-pelan." – English phrases with Indonesian affixes.

Interaction and Response

Insertion is nearly exclusive (15 out of 17 utterances). These consist of short pragmatic markers used to maintain conversational flow, express agreement, or signal topic transitions. Examples:

- 1) "**Exactly.**" – Single English word as a response.
- 2) "**Maybe, iya.**" – English adverb inserted before Indonesian confirmation.
- 3) "**So, aku mikir...**" – English discourse marker so at the beginning of an Indonesian sentence.

The findings demonstrate that code-mixing in Agatha Chelsea's speech is systematic and context-dependent. The dominance of insertion across all contexts aligns with previous studies in Indonesian digital contexts (Situmorang et al., 2023; Saragih et al., 2025), confirming that Indonesian remains the matrix language while English contributes lexical content. The presence of morphological integration (e.g., di-replace, ngetrigger, personality-nya) further indicates that inserted elements are not merely borrowed but actively integrated into Indonesian morphological systems, demonstrating a high level of bilingual competence.

Alternation is most prominent in interpersonal and relational contexts, where it functions as a pragmatic device to mark shifts in perspective or to distinguish between general statements and subjective evaluations. This pattern supports the view that alternation allows speakers to organize information and signal stance changes (Wardhaugh & Fuller, 2015). In the context of discussing relationships, the switch to English often accompanies personal reflections, suggesting that English serves as a linguistic resource for subjective or affective meaning.

Congruent lexicalization is concentrated in academic, scientific, and professional contexts. This finding indicates that when conveying complex ideas, the speaker draws on the shared syntactic structures of Indonesian and English to maintain conceptual fluency and precision. The use of congruent lexicalization reflects a deeper level of bilingual integration, where the speaker does not simply insert English elements but actively blends both languages within a unified grammatical frame. This pattern is particularly evident in explanations of psychological and neuroscientific concepts, as well as in discussions of acting and creative processes.

The variation in code-mixing patterns across contexts confirms that language use is shaped by situational factors. In personal narratives, the speaker maintains stable Indonesian structures with minimal mixing, prioritizing clarity and linearity. In academic and professional discourse, the speaker employs more complex integration to achieve precision and fluency. In interpersonal interactions, alternation serves to manage perspective and emotional expression. This contextual sensitivity reflects the speaker's ability to adapt her linguistic resources to the demands of different communicative situations.

This study contributes to the field of sociolinguistics by demonstrating that Muysken's (2000) typology can be effectively applied to natural spoken data in digital media contexts. Moreover, the integration of structural classification with contextual analysis offers a more comprehensive understanding of code-mixing as a strategic communicative resource. The findings affirm that code-mixing is not a sign of linguistic deficiency but rather a sophisticated bilingual practice that enables speakers to navigate complex communicative demands.

This research provides significant theoretical implications for the development of sociolinguistic studies, particularly in understanding code-mixing as a systematic and context-dependent phenomenon in digital communication. The findings indicate that code-mixing in podcast discourse cannot be treated merely as random borrowing or a sign of linguistic deficiency, but rather as a strategic bilingual practice shaped by conversational context. This study affirms that Muysken's (2000) structural typology remains highly applicable to natural spoken data in digital media, provided that situational factors such as topic, participant roles, and communicative goals are taken into account. By demonstrating that insertion, alternation, and congruent lexicalization each serve distinct pragmatic functions across different contexts, this research extends the original typology beyond structural classification into a discourse-functional framework.

The research results show that code-mixing functions not only as a tool for lexical gap-filling but also as a pragmatic resource to signal perspective shifts, manage interpersonal relationships, and maintain conceptual fluency. In academic and scientific discourse, congruent lexicalization enables precise communication of technical concepts without sacrificing grammatical coherence. In interpersonal and relational contexts, alternation serves to distinguish between general statements and subjective reflections, allowing the speaker to mark stance and emotional involvement. In personal narratives and daily life discussions, insertion predominates as a means of maintaining narrative flow while incorporating commonly used English expressions. These findings demonstrate that code-mixing is a flexible and intentional communicative strategy rather than an arbitrary or deficient language practice.

In addition, this study has implications for the study of digital media and popular culture. The podcast analyzed shows that public figures like Agatha Chelsea actively shape bilingual discourse practices, reflecting how English has become deeply integrated into the linguistic repertoire of urban, educated Indonesians. The presence of morphological integration (e.g., di-replace, ngetrigger, personality-nya) indicates that English elements are not merely quoted but structurally adapted into Indonesian, pointing to ongoing language contact and change. Thus, podcast discourse can be understood not only as entertainment content but also as a cultural artifact that represents contemporary bilingual identity, social values, and evolving communication patterns in Indonesian society. For content creators and educators, these findings suggest that code-mixing can be a strategic resource for enhancing audience engagement, building rapport, and expressing authenticity, rather than something to be corrected or avoided.

Methodologically, this study shows that sociolinguistic analysis of code-mixing in digital media needs to consider both structural classification and conversational context. Future research should move beyond simple frequency counts and adopt a more integrated approach that examines how topic domain, participant roles, and communicative intentions shape code-mixing patterns. This implication opens up opportunities for further research that combines code-mixing typology with discourse analysis, conversation analysis, or audience reception studies. Researchers are also encouraged to explore multimodal elements, such as intonation, pauses, gestures, and facial expressions, which may co-occur with code-mixing to produce meaning in digital communication. With this approach, future studies can delve deeper into how code-mixing is produced, interpreted, and negotiated by both speakers and listeners in the ever-evolving digital communication ecosystem.

4. Conclusion

Based on the findings and discussion, it can be concluded that the utterances of Agatha Chelsea on Raditya Dika's YouTube podcast contain three types of code-mixing based on Muysken's (2000) typology: insertion, alternation, and congruent lexicalization. Among these, insertion is the most dominant type, accounting for 67.6% of the data (117 out of 173 utterances), which indicates that Indonesian functions as the matrix language while English elements are embedded within Indonesian grammatical structures. Alternation appears in 15.0% of the data (26 utterances) and is characterized by switching between Indonesian and English at clause boundaries, whereas congruent lexicalization constitutes 17.4% of the data (30 utterances), reflecting a higher level of structural integration in which both languages are blended within a shared grammatical frame.

In terms of contextual patterns, code-mixing varies systematically across five conversational contexts: in academic and scientific discourse, insertion and congruent lexicalization are employed to maintain terminological accuracy when discussing complex topics; in personal narrative and daily life, insertion predominates as the speaker maintains relatively stable Indonesian structures; in interpersonal and relational discourse, alternation occurs most frequently as a strategy to signal shifts in perspective and distinguish between general statements and subjective reflections; in professional and creative practice, congruent lexicalization enables the expression of complex experiences with greater fluency; and in interaction and response, insertion appears almost exclusively in short pragmatic markers that help maintain conversational flow.

Furthermore, this study demonstrates that code-mixing in Agatha Chelsea's speech is not a random phenomenon but a systematic and context-dependent communicative strategy, as evidenced by the presence of morphological integration such as *di-replace*, *ngetrigger*, and *personality-nya*, which shows that English elements are actively adapted into Indonesian morphological systems and reflect a high level of bilingual competence. Therefore, podcast discourse can be understood as a site where strategic bilingual practices are naturally performed, and where code-mixing serves as a

linguistic resource for achieving clarity, expressing identity, and navigating complex communicative demands in contemporary media contexts.

References

- Aditiawarman, M., Dewirahmadanirwati, D., & Ulya, R. H. (2025). Sociolinguistic Insights into Youth Language Phenomena: Patterns and Influences in South Jakarta through the Lens of #BeritaAkhirPekan Podcast. *AL-ISHLAH: Jurnal Pendidikan*, 17(1), 707–719. <https://doi.org/10.35445/alishlah.v17i1.6843>
- Androutsopoulos, J. (2014). *Mediatization and Sociolinguistic Change*. Walter de Gruyter.
- Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (Third Edit). Sage Publications. <https://doi.org/10.1016/j.math.2010.09.003>
- Hermawan, P. F., Sholihah, N. A., & Sarosa, T. (2025). Analysing Bilingualism Utterances on Agatha Chelsea and Jerome Polin ' s Podcast. *LINGPOET (Journal of Linguistics and Literary Research)*, 06(03), 220–228.
- Muysken, P. (2000). *Bilingual Speech: A Typology of Code-Mixing*. Cambridge University Press.
- Safitri, M. (2021). *A sociolinguistics analysis of code-switching and code-mixing used by netizen comments in Puella Podcast* (p. 6). Universitas Muhammadiyah Surakarta.
- Saragih, E., Pasaribu, T. K., & Sihite, J. E. (2025). An Analysis of Maudy Ayunda ' s Code-Mixing and Code- Switching o n The Merry Riana ' s Podcast. *PROJECT (Professional Journal of English Education)*, 8(2), 368–378.
- Sitinjak, E. R., Sinambela, E., Melysa, C., & Sitompul, E. (2020). An Analysis of Code Switching and Code Mixing on Teacher's Function of Language Utterances. *Cetta: Jurnal Ilmu Pendidikan*, 3(3), 546–555.
- Situmorang, K., Pasaribu, T. K., & Sembiring, R. A. (2023). An Analysis of Code Mixing and Code Switching on Najwa Shihab Podcast "Susahnya Jadi Perempuan" (Catatan Najwa). *ELT Worldwide: Journal of English Language Teaching*, 10(2), 356–366. <https://doi.org/http://dx.doi.org/10.26858/eltww.v10i2.53616>
- Unit Survey APJII. (2025). Survei Penetrasi Internet dan Perilaku Penggunaan Internet Sebagai perwakilan Pengurus APJII, kami dengan bangga mempersembahkan Profil Internet Indonesia 2025. *Asosiasi Penyelenggara Jasa Internet Indonesia*, 15. <https://survei.apjii.or.id/>
- Wardhaugh, R., & Fuller. M, J. (2015). *An Introduction To Sociolinguistics* (7th ed.). Wiley Blackwell.