Journal of Language
Volume 8 Number 1, pp 183-196
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i1.13287

Code Mixing on Beby Tsabina TV: A Sociolinguistics Study of YouTube Content

Enny Nurhaliza[1], M. Manugeren[2], Idawati Situmorang[3]
[1,2]Faculty of Literature, Universitas Islam Sumatera Utara
Medan, Indonesia
[3]AMIK Parbina Nusantara, Pematang Siantar, Sumatera Utara, Indonesia
E-mail: eninurhaliza@gmail.com

Article Info
Article history:
Received: April 08, 2026
Accepted: April 29, 2026
Published: May 11, 2026

Keywords:
alternation;
bilingualism;
code mixing;
congruent lexicalization;
insertion

Abstract
This study analyses code-mixing in the YouTube content of Beby Tsabina, an Indonesian actress and content creator on her channel *Beby Tsabina TV*. Beby frequently mixes English and Indonesian in her videos, presenting a rich case of bilingual language use. Using a descriptive qualitative method, this research identifies types of code-mixing based on Marzona's (2000) theory and describes their functions based on Marzona's (2005) framework across three selected videos. The analysis reveals 14 instances of congruent lexicalization, 45 instances of insertion, and 47 instances of alternation. Functionally, the data show 4 instances of greetings, 31 of storytelling, 11 of expressing opinions, 26 of giving instructions, and 34 of conveying emotions. The findings indicate that alternation is the dominant type of code-mixing, while conveying emotions is the most frequent function. This study contributes to sociolinguistics and digital discourse research by providing empirical evidence of how Indonesian YouTubers strategically employ code-mixing to enhance engagement, express identity, and navigate bilingual communication in online media.

Corresponding Author:
Enny Nurhaliza
Faculty of Literature, Universitas Islam Sumatera Utara, Medan, Indonesia
E-mail: eninurhaliza@gmail.com

1. Introduction

Sociolinguistics is a branch of linguistics that specifically examines the use of language in society. According to Trudgill (1974), as cited in Saputra et al. (2019), "Sociolinguistics is that part of linguistics which is concerned with language as a social and cultural phenomenon." Sociolinguistics examines language functions, varieties, language change, language planning, and people's attitudes toward language use and its users. Within sociolinguistic phenomena, some people use more than two languages. For example, Indonesia has hundreds of different languages due to its ethnic diversity, where groups such as the Javanese, Malay, Batak, and Sundanese possess their own distinct languages. This diversity makes it possible for Indonesian people to speak more than one or two languages. In addition, this phenomenon is also known as bilingualism and multilingualism.

Bilingualism refers to the ability to speak two languages, while multilingualism describes the capacity to use more than two languages (Sukrisna, 2019). This linguistic phenomenon is prevalent worldwide, including in Indonesia, where ethnic diversity naturally fosters a polyglot society. Many Indonesians, particularly within the Gen Z generation, demonstrate proficiency in multiple languages. Furthermore, these individuals frequently engage in code-mixing, a practice that involves the fluid transition between different languages. When individuals systematically combine two or more languages or codes within a single discourse, this practice is formally recognized as code-mixing (Safitri, Harida, & Hamka, 2017).

According to Hudson (1990), as cited by Dewi (2021), code-mixing occurs when bilingual individuals switch from one language to another depending on the shifting social situation. In this process, speakers may incorporate specific words from one language, transition into another, and subsequently return to the primary language. As Muysken in Esnita et al. (2022) suggests, this phenomenon can occur within a single utterance involving two or more languages or language varieties. In the contemporary era, code-mixing is ubiquitous in daily communication across both formal and informal contexts. In this digital age, the practice has become especially prominent on social media platforms such as Facebook, X (formerly Twitter), Instagram, and TikTok.

YouTube, in particular, has emerged as a significant hub for this linguistic trend, as many content creators utilize mixed languages to engage their audiences. An increasing number of individuals globally leverage the internet to convey their opinions, emotions, and artistic expressions through YouTube. This platform's significance is underscored by its status as one of the most visited websites globally, trailing only Facebook, Instagram, and Twitter (Huda, 2017, as cited in Wulandari, Valiantien, & Asanti, 2021). Consequently, YouTube serves as a rich digital landscape for observing the evolution of modern sociolinguistic practices.

Despite growing research on code-mixing in Indonesian YouTube content, previous studies have predominantly focused on established YouTubers with large followings without systematically comparing how different types of content (vlogs, challenges, daily routines) influence code-mixing patterns. Furthermore, existing research has primarily analysed code-mixing in scripted or semi-scripted content, leaving unscripted, spontaneous speech—which better reflects natural bilingual behaviour—underexplored. Additionally, limited attention has been paid to the specific functions of code-mixing in digital discourse, particularly how content creators use language mixing to build intimacy with audiences, express authenticity, or manage face in online interactions.

Beby Tsabina is a significant subject for this study for three reasons. First, as an Indonesian actress, model, and singer with over 1.6 million subscribers on her YouTube channel *Beby Tsabina TV*, she represents a successful crossover from traditional media to digital platforms, making her linguistic choices potentially influential for her large, predominantly young audience. Second, her content spans diverse genres—including daily vlogs, challenges, and behind-the-scenes footage—providing a varied corpus for analysing how code-mixing functions across different communicative contexts. Third, as a Gen Z content creator who frequently uses both English and Indonesian in unscripted, spontaneous speech, Beby offers a naturalistic data source for observing authentic bilingual practices in contemporary Indonesian digital media, distinguishing her from creators who rely on heavily scripted content.

The writers analyse three videos from the *Beby Tsabina TV* channel selected through purposive sampling. These specific videos were chosen because they cover a wide range of conditions, ensuring both the consistency of language use and contextual relevance for the study. To investigate the code-mixing that occurs in Beby Tsabina's utterances, the writers employ a qualitative descriptive method. This research aims to provide language users with a clear representation of how individuals who speak more than two languages utilize their linguistic repertoire and to explore the broader language phenomena inherent in bilingualism.

Accordingly, this research addresses the following research questions:

1. What types of code-mixing are used in Beby Tsabina's YouTube videos based on Marzona's (2000) classification?
2. What communicative functions do these code-mixing instances serve based on Marzona's (2005) framework?

Journal of Language
Volume 8 Number 1, pp 183-196
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i1.13287

Therefore, based on this background, the writers are motivated to conduct a study entitled *Code-Mixing Used by Beby Tsabina on the Beby Tsabina TV YouTube Channel*.

2. Literature Review

2.1 Sociolinguistics

Humans are inherently social creatures who rely on one another to support and satisfy their various needs. As a result, individuals are deeply integrated into society. Language and society are inseparable, as humans must communicate with one another to convey messages, feelings, and ideas. According to Kridalaksana and Djoko Kentjono, as cited in Naibaho (2020), language is defined as an arbitrary system of sound symbols used by members of social groups to cooperate, communicate, and identify themselves. The scholarly study that examines this intricate relationship between language and society is known as sociolinguistics.

Sociolinguistics explores the connection between language and the social fabric, focusing on how individuals utilize language across different social situations. As an aspect of macrolinguistics, sociolinguistics represents the intersection of two distinct fields: sociology and linguistics. Despite their different origins, they maintain a symbiotic relationship. However, it is important to note that the primary focus of sociolinguistics is not sociology itself, but rather the study of language within a social context.

According to Sumarsono (2010), the term sociolinguistics is derived from two words: "socio," meaning society, and "linguistics," referring to the scientific study of language. Sociolinguistics also serves as a vital resource for determining the most effective approach to communicating with others in various settings. Furthermore, the multilingual landscape of a country is often shaped by a variety of complex factors, including its history, geographical environment, and cultural evolution.

Yasemin (2013), as cited in Wilymafidini (2021), defines sociolinguistics as the science that investigates the aims and functions of language within society. This field attempts to explain how language varies across different contexts and geographical borders, as well as how individuals from one context communicate with those from others, such as interactions between non-native speakers or between native and non-native speakers. According to Spolsky (2010), sociolinguistics is the study of the link between language and society, focusing on language variation and attitudes toward language use. This perspective is supported by Hudson (1996), who defines the field as the study of the relationship between language and social factors, including social class, age, gender, and ethnicity. Furthermore, Trudgill (1974) emphasizes that sociolinguistics is the branch of linguistics concerned with language as a complex social and cultural phenomenon.

2.2 Bilingualism

In general, individuals who can express themselves in more than one language are referred to as bilinguals, and the phenomenon of becoming bilingual is known as bilingualism. Bilingualism is defined as the ability to speak and use two languages proficiently (Chaer & Agustina, 1995, as cited in Naibaho, 2020). However, this definition has been debated, as scholars hold different views on what constitutes bilingualism. This definition aligns with the classical perspective provided by Bloomfield (1933), who describes bilingualism as "native-like" control of two languages. Bloomfield's rigorous standard has been critiqued as overly exclusive, as it would exclude many functional bilinguals who use two languages effectively in daily life without achieving native-like fluency in both. Because the term "native-like" can be perceived as an overly rigorous standard, other experts have formulated more practical definitions. For instance, Titone (1972), as quoted by Hamers and

Blanc (2004), claims that bilingualism is an individual's capacity to speak a second language by adhering to its specific concepts and structures rather than merely paraphrasing their mother tongue.

Furthermore, Spolsky (1998) defines a bilingual as a person who possesses some functional ability in a second language, noting that this proficiency may vary significantly from one individual to another. Spolsky's inclusive definition is most appropriate for this study, as Beby Tsabina demonstrates functional ability in both Indonesian and English through her frequent and fluid mixing of the two languages in unscripted YouTube content, even if she may not meet Bloomfield's "native-like" standard. Nevertheless, Steinberg et al. (2001) clarify that bilingualism is distinct from bidialectalism, which refers to a person's ability to use different dialects within a single language. Therefore, to be considered bilingual, a person must be capable of expressing themselves in two or more distinct languages.

Regarding the social context, Hamers and Blanc (2004) define a bilingual community as a linguistic environment where several individuals are bilingual and two languages are in constant contact, often resulting in the use of both languages within the same interaction. Similarly, Fishman (1971), as quoted by Platt and Platt (1975), suggests that bilingualism typically emerges in complex speech communities where individuals become aware of and acquire multiple languages. Gumperz (1982) further observes that bilingual individuals often utilize specific idioms for in-group communication while using a common language for interaction with outsiders.

2.3 Code Mixing

One significant phenomenon that occurs within the field of sociolinguistics in bilingual or multilingual societies is code-mixing. The occurrences of code-mixing and code-switching have long intrigued scholars who examine the various triggers behind these linguistic behaviours (Muysken, 2000, as cited in Wulandari et al., 2021). A key debate in the literature concerns whether code-mixing and code-switching are distinct phenomena or overlapping categories. Some linguists treat them as interchangeable, while others distinguish them based on syntactic boundaries. Code-mixing specifically occurs when bilingual individuals combine words, sentences, or clauses from one language into another within a single discourse. Individuals proficient in more than one language are increasingly likely to employ code-mixing during everyday communication.

In a multilingual society, speakers frequently integrate one language with another, a process widely recognized as code-mixing. According to Suwito (2004), code-mixing involves the usage of two or more languages by injecting portions of one into the other. Furthermore, Trudgill (1992) describes code-mixing as a process where speakers transition between languages, even incorporating words and phrases that may be phonetically challenging in the primary language being used.

As stated by Hudson (1996), code-mixing occurs when a bilingual person who is fluent in communicating with other fluent bilinguals changes the language without necessarily altering the situational conditions. Based on the aforementioned theories, it can be concluded that code-mixing manifests when an individual utilizes two or more languages within the same speech act. Additionally, code-mixing is apparent in both spoken and written forms. This phenomenon is frequently identified in written works such as newspapers, novels, periodicals, academic articles, social media content, and film scripts.

2.3.1 Types of Code Mixing

Three stages have been identified in the typology of code-mixing by Muysken (2000): congruent lexicalization, alternation, and insertion. A critical issue in the literature is whether Marzona (2000, 2005, 2017) offers a distinct framework or merely adapts Muysken's

Journal of Language
Volume 8 Number 1, pp 183-196
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i1.13287

typology. After examination, this study follows the position that Marzona's typology is fundamentally an adaptation of Muysken's work for the Indonesian context. Marzona adopts the same three categories—insertion, alternation, and congruent lexicalization—with almost identical definitions. This study follows Marzona's adapted framework because it has been widely applied in Indonesian sociolinguistic research, providing a consistent basis for comparison with previous studies on Indonesian YouTube content (Rimadhani et al., 2022; Wulandari et al., 2021). The following sections present the three types as defined by Muysken (2000) and operationalized in Marzona's framework.

1. Congruent Lexicalization

Congruent lexicalization is identified as the first type of code-mixing. This phenomenon occurs when the two languages being mixed possess similar or compatible grammatical structures, thereby allowing for a seamless blending of elements from both languages within a single sentence. In this process, the speaker operates within a shared grammatical framework that can be lexically filled with items from either language.

A clear example of congruent lexicalization in an Indonesian-English context is as follows: *"I love makan di sini karena;* the food is so delicious." This sentence can be translated as "I love eating here because the food is so delicious." In this instance, the statement combines English and Indonesian within a single structure that remains grammatically consistent across both languages. The underlined portions of the sentence demonstrate how the mix of languages occurs effortlessly because the underlying structures are almost identical, allowing lexical items to be swapped without disrupting the syntactic flow.

2. Insertion

Insertion is identified as the second type of code-mixing. The concept of insertion corresponds to the term "transference" and "embedding" as introduced by Myers-Scotton (1993). This phenomenon occurs when a specific lexical item from one language is incorporated or "embedded" into the structure of another language. Insertion is often compared to spontaneous lexical borrowing because it is typically limited to a single lexical unit, such as a noun or a verb.

An example of insertion within an Indonesian-English context is as follows: "*Hari ini kita mau* lunch *bareng di kafe dekat sini*." (Today we want to have lunch together at a cafe near here). In this bilingual sentence, Indonesian serves as the matrix language or the primary structure, while an English lexical item is inserted into the statement. The specific word being utilized or inserted in this instance is "lunch," which functions as a noun within the otherwise Indonesian sentence structure.

3. Alternation

Alternation is identified as the final type of code-mixing. This process involves a transition where the structures of two languages alternate distinctly at both the lexical and grammatical levels. Unlike insertion, where one language remains dominant, alternation occurs when two separate grammatical systems are used sequentially. This process can manifest at various linguistic levels, including the replacement of syllables, phrases, or entire clauses with elements from the other language.

A clear example of alternation within an English-Indonesian context is as follows: "*She doesn't like it, tapi dia harus makan*." (She doesn't like it, but she has to eat). This sentence demonstrates a simultaneous mixing and switching of codes within a single

utterance. The first part of the sentence utilizes English grammatical structures, but then transitions into Indonesian starting from the conjunction "*tapi*" (but). Consequently, it can be concluded that this sentence represents the process of alternation, as it showcases a balanced shift between two independent linguistic systems.

2.3.2 Functions of Code-Mixing

The application of code-mixing serves various communicative purposes. While Marzona (2017) identifies five functions—greetings, storytelling, expressing opinions, giving instructions, and conveying emotions—these functions are not mutually exclusive. For example, a YouTuber might mix languages to both tell a story and convey emotion simultaneously, and the functions may be platform-specific. On YouTube, code-mixing for storytelling and emotional expression may serve to create intimacy with viewers, a phenomenon less common in formal written discourse. The following are the five functions as defined by Marzona (2017):

A. Greeting

Greeting is a conventional way of acknowledging or saying hello as a friendly expression when meeting others. In the context of code-mixing, greeting involves utilizing words or phrases from different languages to address individuals in a distinctive manner. This practice may occur due to cultural inclusiveness, a desire for emphasis, or as a stylistic choice.

Example: "Long time no see, bro! *Gimana kabarnya*?"

Analysis: This example demonstrates a mix where the first part of the sentence is in English, followed by Indonesian in the latter part.

B. Storytelling

Storytelling is typically employed to narrate experiences and events by blending two or more languages to make the narrative more vivid and engaging.

Example: "*Jadi, waktu itu aku lagi jalan ke Bukit Lawang, terus* suddenly *ketemu sama bule* from Aussie. Dia ngajakin ngobrol abis itu kita tukeran nomor, haha. Bener-bener *fun experience* banget sih."

Analysis: Based on the example, the speaker utilizes Indonesian as the base language while inserting English phrases to provide more descriptive detail to the story.

C. Expressing Opinions

Expressing an opinion occurs when an individual uses a mixture of codes to convey their personal point of view. Opinions represent beliefs that may or may not be supported by empirical evidence but reflect the speaker's subjective stance.

Example: "*Menurutku, plotnya cukup menarik, tapi* ending-*nya agak mudah ditebak sih. Tapi overall, worth it buat ditonton."*

Analysis: In this instance, the speaker primarily uses Indonesian but incorporates English evaluative adjectives to express their judgment of the film or story.

D. Giving Instructions

Giving instructions involves providing guidance, directions, or advice using a mixture of languages.

Example: "*Mungkin kita bisa coba makan di resto yang lagi* hits *sekarang, atau kalau suka alam,* hiking *di gunung juga seru kok*. What do you think?"

Analysis: The speaker uses Indonesian to offer suggestions and then switches to an English rhetorical question to seek the interlocutor's agreement or opinion.

Journal of Language
Volume 8 Number 1, pp 183-196
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i1.13287

E. Conveying Emotions

Conveying emotion refers to the expression of one's internal feelings through communication. Within code-mixing, language mixtures are used to strengthen emotional nuances or subjective sentiments.

Example: "I'm really happy *lihat kamu hari ini*!"

Analysis: In this statement, the speaker uses an English emotive phrase at the beginning of the sentence, followed by Indonesian to provide context for their happiness.

2.4 Youtube

In general, social media serves as a multifaceted tool for information retrieval, entertainment, commerce, and communication. YouTube stands as the second most popular and frequently utilized social media platform globally, boasting over 2.49 billion users as of 2023. It functions as an expansive database of popular video content, providing a vast array of informative resources. The platform allows users to search for specific information or engage with video content directly. Furthermore, any individual can establish a personal YouTube account to upload content across various genres, including music, film, comedy, horror, travel vlogs, educational courses, tutorials, and podcasts. Those who consistently upload videos to their accounts are commonly referred to as YouTubers.

Moreover, YouTube is recognized as the world's second-largest search engine and the second-most visited website. Compared to other platforms, YouTube is often preferred because its primary medium is video. This format is increasingly favored by modern consumers as it allows for the acquisition of significant information within a shorter timeframe, thereby optimizing efficiency. Based on these developments, YouTube has cemented its position as a dominant social media giant. This global trend is mirrored in Indonesia, which recorded 139 million active YouTube users as of October 2023, making it a fertile ground for sociolinguistic research.

2.5 Beby Tsabina

Beby Tsabina is a prominent Indonesian actress, model, and singer of Acehnese descent. Born in Banda Aceh on October 27, 2002, she began her professional career in the entertainment industry in 2015, gaining initial recognition through her role in the television series *Anak Jalanan*. Beyond her acting career, she is highly active on digital platforms, specifically through her YouTube channel, *Beby Tsabina TV*, which has amassed over 1.6 million subscribers. Her channel features a diverse range of content, including daily vlogs, interactive challenges, and musical covers.

Several of the most popular videos on her channel underscore her significant digital reach, such as *Beby Tsabina's First Vlog* with 10 million views, the *24-Hour No Makeup Challenge* with 7 million views, and her cover of Rizky Febian's *Bersamamu* with 6 million views. She utilizes YouTube primarily as a medium to interact with her fanbase, frequently sharing personal stories and behind-the-scenes experiences from her life as an actress. The success of her channel demonstrates the viability of crossover success for traditional media figures within the digital landscape of YouTube in Indonesia.

For the purpose of this study, the writers selected specific videos from Beby's channel for detailed linguistic analysis. The focus of the investigation is to identify the various types and functions of code-mixing present within her utterances. These particular videos were chosen because they contain a high frequency of English lexical items integrated into her Indonesian speech, providing a rich data set for observing contemporary code-mixing phenomena.

2.6 Review of Related Findings

In this section, the writers examine previous research efforts, including two distinct studies that serve as reference points for the current investigation. These studies are compared critically to identify what has been done and what remains underexplored.

The first study was conducted by Zul Astri and Al Fian (2020), entitled *"The Sociolinguistics Study on the Use of Code-Mixing in Gita Savitri Devi's YouTube Channel Video."* Similar to the present study, their research focuses on Indonesian-English code-mixing using a descriptive-qualitative method. However, the primary difference lies in the theoretical framework. Zul Astri and Al Fian utilized Hoffman's theory to categorize the types of code-mixing into three forms: intra-sentential, intra-lexical, and those involving changes in pronunciation. Hoffman's typology focuses on syntactic placement (where mixing occurs), whereas the present study adopts Muysken/Marzona's typology, which distinguishes mixing based on grammatical relationship (how languages relate structurally). Their findings identified 52 instances of code-mixing across categories such as words, phrases, sentences, prefixes, and suffixes, with intra-sentential code-mixing being the most frequent and pronunciation changes the least. In contrast, the current study employs Marzona's theory to analyze both the types and functions of code-mixing.

The second reference is a study by Rimadhani et al. (2022), titled *"Phenomena of Code-Mixing on YouTube Videos: A Case Study of Pita's Life YouTube Account."* This descriptive qualitative research analyzed nine videos uploaded in 2018, focusing on Indonesian-English code-mixing. This study aligns more closely with the current research as it also utilizes Marzona's theory. However, there are important differences. Rimadhani et al. identified three types of code-mixing (insertion, alternation, and congruent lexicalization) and investigated the motives behind these linguistic choices, classifying them into the fulfilment of needs and the pursuit of prestige. Their results revealed 18 cases of insertion, 2 cases of alternation, and 4 cases of congruent lexicalization, with motives identified in 5 data points (3 for fulfilling needs, 2 for prestige).

The research gap this study addresses is twofold. First, existing studies have focused on established YouTubers without systematically comparing how different types of content influence code-mixing patterns. Second, previous research has primarily analysed motives (why speakers mix languages at a psychological level) rather than communicative functions (what purposes mixing serves in discourse). This study extends previous research by applying Marzona's functional framework to a highly popular YouTuber (Beby Tsabina, with over 1.6 million subscribers) and by analysing both structural types and communicative functions in spontaneous, unscripted speech.

3. Research Method

This research employs a qualitative descriptive design to analyze the code-mixing utilized by Beby Tsabina. According to Moleong (2018), qualitative research seeks to understand phenomena related to participants' conditions, including behavior, perceptions, motivations, and actions. Nassaji (2015) explains that descriptive research aims to categorize and outline the phenomena under study. To achieve this, the writers apply passive participant observation, watching Beby Tsabina's YouTube videos and recording instances of code-mixing using Marzona's (2017) theory.

Three videos from the *Beby Tsabina TV* YouTube channel were selected using purposive sampling based on four criteria: (1) Beby Tsabina as the primary speaker, (2) diverse content genres (daily vlog, weekly recap, challenge), (3) publication within the same year, and (4) high frequency of English-Indonesian mixing. The selected videos are: "Beby's

Journal of Language
Volume 8 Number 1, pp 183-196
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i1.13287

VLOG #138" (23:47), "Beby's VLOG #139" (21:15), and a third video (19:32), totalling approximately 64 minutes of analysed content.

The data collection process follows a documentation technique. The systematic procedure involves watching the selected videos, transcribing utterances, identifying code-mixing instances, classifying data based on Marzona's types and functions, and interpreting the findings.

To ensure reliability, intra-rater reliability was conducted by re-analysing the same data after a two-week interval (92% agreement). Inter-rater reliability was also conducted by having a second coder independently code 20% of the data (87% agreement), following Creswell (2014) and Miles, Huberman, and Saldaña (2014).

Ethically, all videos are publicly accessible on YouTube without privacy restrictions. This study is purely observational, non-commercial, and does not interact with or harm the content creator. All data are properly attributed to Beby Tsabina as the original source, following the principles of the Association of Internet Researchers (Markham & Buchanan, 2012).

4. Discussion

In this part, the writers focus the analysis on the purpose that has been mentioned in the objective of the study, which is to identify the type of code-mixing and describe the function of code mixing seen in Beby Tsabina's utterances in her YouTube videos. This study uses transcript data taken randomly to identify the type and describe the function of code-mixing that occurs in the video.

4.1 Type of Code Mixing

There are three types of code mixes that occur in Beby Tsabina's speech in her three videos, the writers use Marzona's theory to identify these types of code mixes. The following below is a more detailed explanation.

1. Congruent Lexicalization

a. (Beby's VLOG #138 A Day in My Life as a Student in Western Michigan University – minute 3:13 - 3:17)
"*Pokoknya hari ini bakal* very productive *dan sangat sibuk banget*"
(anyway today will be very productive and very busy)
Code Mixing: very productive

The utterances above show English words or phrases that have similar or the same equivalents in Indonesian. There are English words "very" and "productive" whose meanings in Indonesian are "*sangat*" and "*produktif*" According to Marzona's theory of code-mixing, this example fits into the category of Congruent Lexicalization, where elements from two languages with similar grammatical structures are combined.

2. Insertion

a. (Beby's VLOG #138 A Day in My Life as a Student in Western Michigan University – minute 2:36 – 2:40)
"*Ini emang sebenarnya* the most wanted *video dari kalian semua*."
(This is actually the most wanted video from all of you)
Code Mixing: the most wanted

The phrase "the most searched video" is inserted into the Indonesian sentence, indicating an English element inserted into the Indonesian sentence structure. This fits into the category of Insertion according to Marzona's theory.

3. Alternation

a. (Beby's VLOG #138 A Day in My Life as a Student in Western Michigan University – minute 17:25 – 17:30)

"Hii im done, *udah selesai kelasnya* for today."

(Hii, I'm done. The class is over for today.)

Code Mixing: Hii I'm done, for today

This utterance shows an alternation between English and Indonesian. The first part "Hii I'm done" is in English, and the second part "udah selesai kelasnya for today" is a mixture, of "for today" in English. This sentence fits into the alternation of Marzona's theory.

The analysis revealed 14 instances of congruent lexicalization, 45 instances of insertion, and 47 instances of alternation. Alternation emerged as the most frequent type, while congruent lexicalization was the least frequent. This dominance of alternation can be explained by several factors. First, YouTube vlogs involve extended narrative monologues where speakers move between topics; alternation allows Beby to signal topic shifts or emotional changes by switching languages. Second, unlike insertion which requires embedding single words into a dominant grammatical frame, alternation does not require either language to be dominant, offering greater flexibility for spontaneous speech. Third, congruent lexicalization requires the two languages to share similar grammatical structures. Indonesian and English differ significantly in syntax (e.g., Indonesian has no verb tense marking while English has complex tense systems), making congruent lexicalization more difficult to produce naturally. This finding supports Muysken's (2000) observation that congruent lexicalization is most common in closely related languages rather than typologically distant pairs like Indonesian-English.

4.2 Function of Code mixing

a) Greetings

"Hi hi welcome back, *di video kali ini aku mau bikin* a week in my life" *(#139 Beby's VLOG: A Week in My Life: Study, Fashion Show & Movie Day – minute 00:36 – 00:42)*

The utterance above shows a greeting in English followed by Indonesian and then English again. Beby Tsabina opens the initial greeting "hi hi welcome back" in her video to greet his audience followed by the sentence "*di video kali ini aku mau bikin*" then "a week in my life" as a context that describes the content of his video.

b) Story Telling

"Im just happy right now, *karna tadi sebelum kelas, aku ketemu sama orang kayak* Korean-american *gitu main* skate, *terus kayak* he ask for my number" (*Beby's VLOG #138 A Day in My life as a Student in Western Michigan University – minute* 17:40 – 18:01*)*

The utterance pad above shows the function of storytelling code mix which uses a mixture of English and Indonesian to convey the story more expressively. In her utterance, Beby tells her personal experience when meeting someone.

c) Expressing Opinion

"*Sekarang aku lagi suka bangett pake ini karna enak banget berasa kayak dipijet gitu bagus juga buat sirkulasi darah* so its really good" (*Beby's VLOG #138 A Day in My life as a Student in Western Michigan University – minute 01:47 – 01:55)*

Based on the above speech showing the function of conveying opinions, Beby mixes Indonesian as the initial part of the sentence followed by English to convey his opinion about the use of guasha which gives a massage sensation.

Journal of Language
Volume 8 Number 1, pp 183-196
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i1.13287

d) Giving Instructions
"*Pertama aku* wash my face *abis itu bilas seperti biasa*" (*Beby's VLOG #138 A Day in My life as a Student in Western Michigan University – minute 01:27 – 01:30)*
The utterance above shows the code-mixing function of giving instructions in which one phrase "wash my face" is inserted in the Indonesian sentence structure. Beby gives instructions about her skincare routine by mixing Indonesian and English to provide a more detailed explanation.

e) Conveying Emotion
"*Katanya di* show *kalo aku* researchnya *banyak dan aku kek beneran tau* about the company it self, so im very happy."*(#139 Beby's VLOG: A Week in My Life: Study, Fashion Show & Movie Day – minute 12:54 – 13:02)*
The utterance above shows the code-mixing function of conveying emotion which is Indonesian and then followed by English. Beby showed a happy feeling with the phrase "so I'm very happy" to indicate her positive feelings because of the recognized effort.

The functional analysis showed that conveying emotions (34 instances) and storytelling (31 instances) were the most frequent functions, while greetings (4 instances) were the least frequent. This distribution reflects the nature of YouTube vlogs as personal, narrative-driven content.

Storytelling is frequent because vlogs are essentially first-person narratives of daily experiences. Beby uses English to highlight exciting moments ("Korean-american," "skate," "ask for my number") and Indonesian to establish conversational intimacy. This pattern illustrates what Myers-Scotton (1993) calls the "markedness model": speakers choose the unmarked language (Indonesian) for neutral narration but switch to the marked language (English) for emphasis, evaluation, or dramatization.

Conveying emotions is equally frequent because vlogs are personal and emotionally expressive. Beby repeatedly uses English phrases ("I'm so happy," "so I'm very happy") to express positive emotions. English serves as a "distancing" mechanism, allowing her to express strong emotions while maintaining a degree of performative control—a pattern documented in digital discourse studies. For negative emotions, she typically stays with Indonesian, suggesting that English functions as a language of positive affect while Indonesian carries heavier emotional weight for vulnerability.

Greetings are rare because YouTube vlogs, unlike live conversations, do not require repeated greetings. Beby typically greets her audience once at the beginning of each video, but the greeting function is often fulfilled through video titles, thumbnails, or non-verbal cues, reducing the need for verbal greetings. This finding highlights that the frequency of communicative functions is genre-dependent, and researchers must interpret functional distributions within their specific media contexts rather than as universal patterns.

4.3 Broader Sociolinguistic Implications

These findings contribute to broader theories in sociolinguistics in three ways. First, they support Muysken's (2000) typology by demonstrating that the three types of code-mixing are empirically distinguishable and occur in different proportions depending on linguistic and contextual factors. The rarity of congruent lexicalization in Indonesian-English mixing confirms Muysken's observation that typological distance between languages shapes code-mixing patterns.

Second, the findings extend Gumperz's (1982) theory of conversational code-switching to digital contexts. Gumperz identified six functions of code-switching (e.g., quotation, addressee specification, interjection). This study's functional categories

(storytelling, conveying emotions, giving instructions) align with Gumperz's framework while adding platform-specific functions (e.g., greeting YouTube audiences, creating vlog narrative structure). This suggests that digital discourse requires an expanded functional typology that accounts for one-to-many, asynchronous communication.

Third, the study contributes to research on language and identity in digital media. Beby's code-mixing patterns construct a hybrid identity: globally connected (through English) yet authentically Indonesian (through Indonesian for personal content). This bilingual identity resonates with an increasingly globalized generation of Indonesians who navigate multiple linguistic and cultural worlds daily. As English continues to spread through digital media, code-mixing like Beby's may become the default for young, urban, educated Indonesians rather than a marked choice.

5. Conclusion

After analyzing with using Marzona's theory, the writers have found that there are three types of code-mixing in the three Beby Tsabina videos mentioned in the scope of this research. The types of code mixing are congruent lexicalization, insertion, and alternation.

Table 1

Frequency Distribution of Code-Mixing Types

Type of Code-Mixing	Frequency	Percentage (%)
Congruent Lexicalization	14	13.2
Insertion	45	42.5
Alternation	47	44.3
Total	**106**	**100**

Note:

N = 106. Percentages are rounded to one decimal place.

Based on the data shown in the table, it can be concluded that there have been several types of code mixing observed in Beby Tsabina's utterances in her three videos on the Beby Tsabina TV channel. In this research series, the writers have found 14 data of congruent lexicalization, 45 data of insertion, and 47 data of alternation. Significantly, alternation has been the most frequently occurring type of code-mixing, while congruent lexicalization has been the least frequently encountered in the utterances.

Furthermore, the writers have also identified the functions of code mixing present in three videos of Beby Tsabina. These include greetings, storytelling, expressing opinions, giving instructions, and conveying emotions, as outlined in Marzona's theory adopted by the writers.

Table 2

Frequency Distribution of Code-Mixing Functions

Function of Code-Mixing	Frequency	Percentage (%)
Greetings	4	3.8
Storytelling	31	29.2
Expressing Opinions	11	10.4
Giving Instructions	26	24.5
Conveying Emotions	34	32.1
Total	**106**	**100**

Note:

N = 106. Percentages are rounded to one decimal place.

Journal of Language
Volume 8 Number 1, pp 183-196
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i1.13287

From the table above, it can be concluded that several instances have shown the functions of code-mixing, namely 4 data of greetings, 31 data of storytelling, 11 data of expressing opinions, 26 data of giving instructions, and 34 data of conveying emotions.

References

Astri, Z. & Fian, A. (2020). The sociolinguistics study on the use of code-mixing in Gita Savitri Devi's YouTube channel video. *Journal of English Language Teaching and Literature (JELTL)*, *3*(2), 154-165.

Bloomfield, L. (1933). *Language*. Henry Holt and Company.

Chaer, A., & Agustina, L. (1995). *Sosiolinguistik: Perkenalan awal*. Rineka Cipta.

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

Dewi, K. T. (2021). Language use: Code mixing, code switching, borrowing, pidginization, and creolization. *Yavana Bhasha: Journal of English Language Education*, *4*(1), 34-44.

Esnita, E., Syafitri, D., Rahayu, S., & Putri, N. (2022). An analysis of code mixing used by social media users in digital communication. *Journal of Language and Literature*, *10*(1), 45-56.

Fishman, J. A. (1971). *Sociolinguistics: A brief introduction*. Newbury House.

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge University Press.

Hamers, J. F., & Blanc, M. H. (2004). *Bilinguality and bilingualism* (2nd ed.). Cambridge University Press.

Hoffman, C. (1991). *An introduction to bilingualism*. Longman.

Huda, M. (2017). *Pengaruh media sosial YouTube terhadap perilaku komunikasi mahasiswa* [Undergraduate thesis]. Universitas Islam Negeri Sunan Kalijaga.

Hudson, R. A. (1996). *Sociolinguistics* (2nd ed.). Cambridge University Press.

Kridalaksana, H. (2008). *Kamus linguistik* (4th ed.). Gramedia Pustaka Utama.

Markham, A., & Buchanan, E. (2012). *Ethical decision-making and Internet research: Recommendations from the AoIR ethics working committee* (Version 2.0). Association of Internet Researchers.

Marzona, Y. (2017). The use of code mixing between Indonesian and English in Indonesian advertisement of Gadis. *Jurnal Ilmiah Langue and Parole*, *1*(1), 238-248.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage Publications.

Moleong, L. J. (2018). *Metodologi penelitian kualitatif*. Remaja Rosdakarya.

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

Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in code-switching*. Clarendon Press.

Naibaho, L. (2020). *Introduction to sociolinguistics*. UKI Press.

Nassaji, H. (2015). Qualitative and descriptive research: Data type versus data analysis. *Language Teaching Research*, *19*(2), 129-132.

Platt, J. T., & Platt, H. K. (1975). *The social significance of speech: An introduction to and workbook in sociolinguistics*. North-Holland Publishing Company.

Rimadhani, N. R., Sari, D. P., Hartanti, M. P., & Wulandari, T. (2022). Code mixing phenomenon in YouTube video: A case study of Pita's Life account. *Ilmu Budaya: Jurnal Bahasa, Sastra, Seni, dan Budaya*, *6*(2), 579-588.

Safitri, M., Harida, E. S., & Hamka. (2017). An analysis of code mixing used by students of TBI in social media (Facebook) at fifth semester of IAIN Padangsidimpuan. *English Education: English Journal for Teaching and Learning*, *5*(2), 52-64.

Saputra, J., Rahmawati, A., & Setiawan, B. (2019). The importance of sociolinguistics in foreign language education: A review of selected paper. *International Journal for Educational and Vocational Studies*, *1*(4), 299-303.

Spolsky, B. (1998). *Sociolinguistics*. Oxford University Press.

Spolsky, B. (2010). *Sociolinguistics*. Oxford University Press.

Steinberg, D. D., Sciarini, N. V., & Sciarini, S. J. (2001). *Psycholinguistics: Language, mind and world* (2nd ed.). Longman.

Sukrisna, A. (2019). *An analysis of code mixing used by English teachers in teaching learning process at Mas Darul Ulum Tanjung Bintang in the academic year of 2018/2019* [Undergraduate thesis]. UIN Raden Intan Lampung.

Sumarsono. (2010). *Sosiolinguistik*. Pustaka Pelajar.

Suwito. (2004). *Sosiolinguistik: Teori dan masalah*. Penerbit Universitas Sebelas Maret.

Titone, R. (1972). *Le bilinguisme: Théorie et pratique*. Didier.

Trudgill, P. (1974). *Sociolinguistics: An introduction*. Penguin Books.

Trudgill, P. (1992). *Introducing language and society*. Penguin Books.

Wilymafidini, O. (2021). Code mixing used by the teachers of LPAM English Course Semurup in the teaching learning process. *Inovish Journal*, *6*(1), 65-74.

Wulandari, T., Valiantien, N. M., & Asanti, C. (2021). Code mixing in Seleb English video content on YouTube. *Ilmu Budaya: Jurnal Bahasa, Sastra, Seni, dan Budaya*, *5*(3), 503-517.

Yasemin, K. (2013). Sociolinguistics and language education. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley-Blackwell.