CHAPTER I

INTRODUCTION

1.1 Background of The Research

A code is a term used to refer to a system of communication that is used among people. This term, in most cases, is also called “language” (Wardhaugh, 2006.). Some people have the ability to communicate in more than one codes or languages. They are referred to as bilingual or multilingual. In oral or written communication, bilingual speakers often take from the other language a word, a phrase, or a sentence for some reasons, and move from one language to another in one single conversation. The first is called code-switching while the latter is code-mixing.

Scotton (1993) described expressions that combine the grammar of two languages without changing the grammar of the first language used, referred to as code-mixing. Code-mixing is defined as the phenomenon of language where people mix more than one language and even their dialect or language varieties when they are speaking. He argues that code-mixing relates to all cases in which two languages' lexical items and grammatical features appear in one sentence.

Hoffman (1992) stated there are several reasons for a bilingual person to mix their language, talking about a particular topic, quoting, showing empathy about something, interjection (using sentence connector), repetition used for clarification, expressing group identity, and clarifying the speech content to the interlocutor The following is an example of code-mixing between English and Persian (Mahootian, 2006):

I'm **shuxi-ing** with you. The meaning is 'I'm **joking** with you.' (Mixing between English and Persian).

When the speaker says the above sentence, the speaker inserted English word **joking** and the speaker adds –ing in this sentence. In the Persian language, there is no present participle form (–ing), but naturally, the context of the sentence is in English, so the speaker adds –ing in the sentence.

'Sometimes I'll start a sentence in English **y termino en espan˜ ol**'
'Sometimes I'll start a sentence in English **and finish it in Spanish.**'
(Switching between English and Spanish).

. We can see the code mixing occurs in the example above, one sentence consists of two languages. When the speaker says the above sentence, the speaker starts the sentence in English and finishes it with the Spanish language.

Code-mixing is often used in any field related to language use, such as on social media platforms. This phenomenon can be found in verbal communication and also be found easily in non-verbal communication in social media. The following are the examples of code-mixing from YouTube:

Picture 1. TUTORIAL MAKEUP WISUDA FLAWLESS TAPI CETAR! (Ari Izam)

Picture 2. MAKE UP TUTORIAL FOR BEGINNER (Makeup Pemula) – Simple dan gampang! (Abel Cantika)

In the pictures above, the content creator uses two languages, Bahasa Indonesian and English, for the content's title. For **picture 1,** Ari Izam uses Indonesian and then switches to English; for **picture 2,** Abel Cantika uses English and switches to Indonesian.

As previously mentioned, the phenomenon of code-mixing is common in daily life and also in social media such as YouTube. YouTube is the largest and most widely used video platform now. The use of YouTube is not only for entertainment, but it is also commonly used in education as a learning support tool, marketing, and so on. We can easily get information through YouTube, such as the latest information about politics, economics, and history. We can get everything through this platform. Several advertisements and YouTube content use language

phenomena such as code-mixing, such as advertisements for food products, drink products, education, life insurance, etc.

The data source for this research is a video from YouTube; the writer took the utterances from six videos. Three of them are YouTubers who have an educational background in English-speaking countries, and the other three are YouTubers who have an education in Indonesia. There are several reasons why this kind of topic was chosen to be the object of this research. The first reason is that the content is easy to find, and in the video, it is easy to find the phenomenon of code-mixing. Mostly, the YouTubers in makeup tutorial videos convey the content in Indonesian. Usually, they use English to mention the product's name and the steps in the makeup tutorial. This research is exciting to study because it shows the difference between the types of code-mixing used by YouTubers from different educational backgrounds and to see the correlation between the use of code-mixing and English educational background.

This phenomenon of code-mixing caught the writer's attention because when people want to change or mix their languages, it means that they are fully fluent in two or more languages. In this case, code-mixing is considered a strategy by content creators/YouTubers to attract their viewer's attention. Sometimes they mix languages with English to show the interlocutor/viewers that they also know another language or because the other language sounds better. Therefore, the pattern or type of code-mixing used by YouTubers is interesting to study and compare because they come from different educational backgrounds. The correlation between code-mixing and English educational background is shaped by

proficiency, identity construction, audience engagement, and digital language trends. Those with better English education are more confident in code-mixing, whereas those with limited English exposure may mix languages differently or less frequently.The writer is curious to see if the English educational background correlates with the use of code-mixing in the video tutorial makeup Consequently, it is important to analyse this phenomenon of code mixing so that readers can distinguish the type of code mixing used.

1.2. Theoretical Frameworks

1.2.1. Sociolinguistic

Sociolinguistics is the study of language in relation to society. Wardhaugh (1998) explains, "A society is a group of people who come together for one or more specific purposes, and language is what is spoken by a particular member of society". One's language is always spoken in the same way, the difference in the language they use for different purposes.

Sociolinguistics not only addresses the use of language but also all issues related to language behavior, such as language attitudes and language speakers. Sociolinguistics "studies the field of language and society and is closely related to the social sciences" Trudgill (2000). Sociolinguistics also speaks of spoken communities, which are "groups of people who interact through language" Bloomfield (1993). He added that there must be a regular relationship between the use of language and social structure.

1.2.2. Bilingualism

In sociolinguistics, we learn about "bilingualism", the capacity of speakers to utilize two dialects with the same capacity. Bilingualism is the "native-like ability to use two languages". However, this definition excludes many people who speak more than one language but are not native-like in one or both. Many people who regularly use two languages may not be native-like in one of their languages. In the United States, immigrants with the highest levels of education are more likely to try to raise their children to be bilingual, in contrast to the language shift observed among less advantaged groups. In all of these cases, the two languages appear to play slightly different roles in the lives of bilinguals, Bloomfield (1933).

Wardhaugh (2010) described bilingual or plurilingual as a speaker who can speak more than one language or can speak in two or more codes and shift their code when speaking. Plurilingual nature causes people to do language phenomena called code-mixing and also code-switching. An interesting example of multilingualism exists among the Tukano people of the northwestern Amazon, on the border between Colombia and Brazil Sorensen (1971). The Tukano are multilingual because men must marry outside their language group; that is, no man can have a wife who speaks his language because this type of marriage is not allowed and would be considered a form of incest. Men choose the women they marry from neighboring tribes that speak other languages. Multilingual or bilingual also produces a phenomenon called code-mixing and code-switching.

Scotton (2006) described bilingualism as a term that refers to speaking one or more languages. Usually, the mother tongue or the speaker's first language is one

of the two languages that make the person bilingual. Being bilingual does not mean full mastery of two languages. Further speakers are rarely equally fluent in two languages. All normal human intelligence speaks at least one language. Speakers are called bilingual when they have acquired or learned to speak or understand at least some expressions that reveal the internal structural relationships of a second language. Being bilingual is associated with the ability to speak two or more languages, not just the ability to read L2 with an existing dictionary. Just saying a few phrases (such as greetings and equivalents like "please" and "thank you") is not enough to qualify as bilingual.

1.2.3. Code-mixing

Aitchison & Wardaugh (1987) stated that when two or more people talk to each other, they use a communication system called a code. Typically, the code will be represented in a single language. When two codes or languages are combined without changing the context, we discuss mixing codes. Mixed codes often appear in sentences.

Usually, people do code-mixing in certain situations, such as in the case of social situations; people tend to mix their code (language) in social situations so that code-mixing can be associated with a certain addressee. People change their language to show that they belong to a certain group. People can change from one language to another language. Holmes (1992) stated that sometimes, they do not even understand a foreign language completely, but they can use certain words and phrases. Switches are often brief and are mainly done for social purposes, namely, to create a sense of solidarity with those who use them.

It is possible to move between clauses. We may deduce that code-switching happens when people switch languages, dialects, or styles based on the claims made by Brown and Attardo (2000). They claim that code-switching occurs when people change language, dialect, or style in one sentence or near pair.

According to Holmes (1992), code-switching happens when people shift from one language to another. As stated above, code-switching can be divided into inter-sentential and intra-sentential. Inter-sentential language switching occurs when two languages change at a sentence or clause boundary. Intra-sentential language switches are those that occur within a clause and involve a phrase, a single word, or morpheme boundaries refers to Mahootian (2006), and some researchers have used the term code-mixing to refer to intrasentential switching and this research conduct refers to code-mixing. Nonetheless, "code-mixing" and "code-switching" relate to the same kinds of language mixing and are used interchangeably in most current literature.

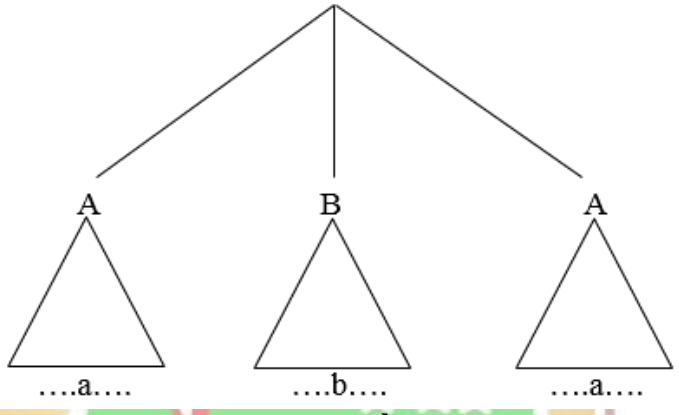
1.2.4. Types of Code-mixing

Hoffman (1991) stated that there are three types of code-mixing based on the syntactic model. These are intra-sentence code-mixing, intra-word code-mixing, and mixing involving changes in pronunciation. Three types of code-mixing mentioned by Muysken (2000) are:

1. **Insertion**

The incorporation of lexical items or entire components of one language into the structure of another language is known as Insertion. It's a kind of

unidirectional language influence, Muysken (2000, p.3). Simply it is a single constituent B (with words b from the same language) inserted into a structure defined by language A (with words a from that language). The following example from Muysken (1987):



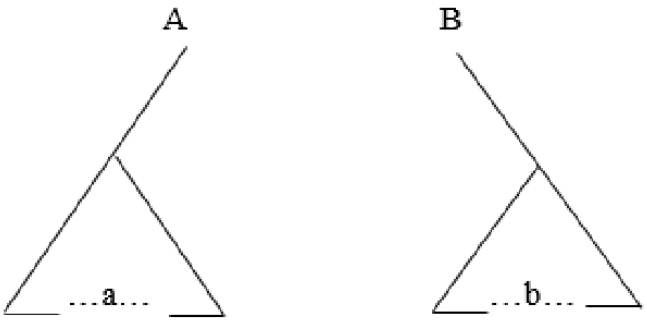
na'iish **crash** la

This is an example of an insertion between Navaho and English. In this utterance, crash is an English verb mixed with the structure of the Navaho language. In this sentence, the English word "**crash.**"

2. Alternation

Alternation is defined as switching between structures from separate languages. The constraint can be a clause or some element, such as a discourse marker or tag form. Bilingual grammar is used autonomously or independently. According to Muysken, there are three points about the syntactic question of alternation, (i) a language A or B structure dominates the sentence containing alternation, or (ii) whether it is the sum of constituents from language A and B or (iii) whether it is simply a non-language-specific node, in which case categorical

equivalence is assumed, Muysken (2002, p.31). Example of alternation (English/Spanish):

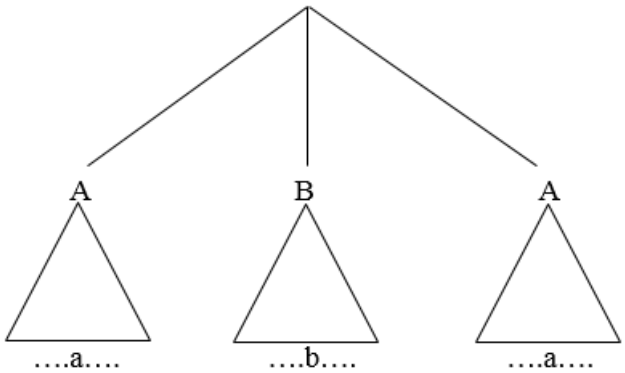


Andale pues and do come again

In the case of alternation, there is a true switch from one language to the other, involving both grammar and lexicon. In this utterance, "**andale pues**" is a Spanish clause mixed with the English phrase "and do come again."

3. Congruent Lexicalization

Congruent lexicalization is when the languages share a grammatical structure, but the vocabulary comes from two or more languages. Muysken (1995) stated that in congruent lexicalization, the syntax of the two languages in the sentence is the same, thus allowing the lexical items from either language to be used. So, it simply mixes the grammatical structure from languages A and B and randomly inserts the words from languages A and B. The following example from Muysken (1987, p.117):



This is an example of congruent lexicalization between Hebrew and English. In this utterance, English and Hebrew share grammatical structure with the form a,b,a. The sentence contains a word from Hebrew in the middle, ending in English.

1.2.5. Language Variation

According to Wardhaugh (2006), language variation is a way of describing certain variations that cause speakers of a particular language to speak a different dialect of that language. Variety is a feature of language, as there are many ways to express the same thing. Variation can include changes in pronunciation, grammar, vocabulary, and even speaking strategies. Language varies according to its uses and users, where it is used, to whom it is used, and who utilizes it.

Labov (1966) described that language variation is systematic rather than random. Individuals use different forms of language depending on their social identity and the context in which they speak. Labov's studies demonstrated that language use often correlates with social factors, such as a person's social status, age, gender, education level, and regional location. The use of a particular language

is strongly influenced by the social dynamics of a given setting, which has been found to impact how people use language significantly. One of the factors is education influences language according to what is learned in school, the effect of education is cumulative. Children's use of linguistic variables is determined by how much schooling they have received, not the general educational milieu of the family, Labov (2001).

Labov has conducted several studies related to the correlation between language style and language variation with social factors. Study of sound change on the island of Martha's Vineyard (Labov 1963). This earlier work tracked the distribution of a particular sound feature as it varied among various occupational, ethnic, and geographic groups of the population. Then study of Taheran in 1978, this study investigated the raising of /a/ to /u/ before nasals, and the use of the third singular possessive /æsh/. Twenty-one subjects were interviewed in Ghazvin. Subjects were distributed across four educational levels,

1.3. Review of Previous Studies

Studies related to code-mixing have been widely carried out in literature, either in written or oral form. This chapter will discuss a previous study of the use of code-mixing in YouTube content.

The first research by Fajrini (2017) aims to discover the types of code-mixing in the host's utterances. She analyzes the types and reasons for the code-mixing that the host utters on "I Look" on NET TV. This research used a theory from Muysken (2000) for the types of code-mixing and the context of ethnography

of communication by Hymes (1972). This research uses a quantitative descriptive method.

This research found three types of code-mixing: insertion, alternation, and congruent lexicalization. It also found 19 data, 15 insertions, 2 alternations, and 2 congruent lexicalizations. The dominant type of code-mixing used by the host is insertion.

Sari and Hartanti (2023) conducted the second research. The study aims to determine the forms of code-mixing and code-switching that influence their utilization and to identify interference in Maudy Ayunda's speech as an effect of the use of code-mixing and code-switching. This research uses a qualitative method. The researcher employed Hoffman's theory (1991).

From this research, the researcher can conclude that Maudy Ayunda more dominantly uses intra-sentential mixing in his video. In some contexts, vocabulary or sentence constructions in one language may be more appropriate or more accessible in conversations in another language. From this research, 176 data were found in total, that is, 109 data on Intra Sentential Code-mixing, 11 on Intra Lexical Code-mixing, and 14 on Involving Change Pronunciation. The researcher discovered that three types of code-mixing, insertion, alternation, and congruent lexicalization, were employed in this novel. Additionally, Maudy Ayunda uses three different sorts of code-switching, emblematic switching 6 times, intra-sentential switching 68 times, and inter-sentential switching 14 times.

The third research was by Situmorang et al. (2023). This research aims to identify the types and how code-mixing and code-switching are used in the podcast. This research uses a qualitative descriptive method. The researcher employed the theory by Muysken (2000). From this research, were found 87 data obtained from this podcast. There are 62 code-mixing data consisting of 46 insertions, 10 alternations, and 6 congruent lexicalizations. Furthermore, 25 code-switching data were found, which consisted of 9 inter-sentential switching, 3 tag switching, and 13 intra-sentential switching.

The fourth research by Yow and Patrycia (2015) examines the relationship between bilingual children's language proficiency and their code-switching behavior. Presented at the 39th Boston University Conference on Language Development, the research challenges the traditional view that code-switching indicates linguistic incompetence. Instead, the findings suggest that higher language competence predicts the frequency and type of code-switching, indicating that code-switching is a marker of linguistic competence rather than deficiency.

The result of this research, bilingual children with stronger language skills were more likely to code-switch, suggesting that code-switching requires cognitive and linguistic ability rather than reflecting a lack of knowledge. Code-switching was used deliberately and appropriately, depending on the conversational context, rather than as a result of struggling with a language, and the study contradicts the "Linguistic Incompetency Hypothesis", which assumes that bilinguals switch languages due to gaps in vocabulary or limited proficiency.

The fifth research by Hutriani (2019) focused on the use of code-mixing by the “Break Out” Music Program hosts. The main focus of this research is the types and reasons for using code-mixing. This research uses the theory by Muysken (2000) to identify the types of code-mixing and Hoffman (1991) to explain the reasons for code-mixing. This research uses a qualitative method.

The result of this research, there are three types of code-mixing insertion, alternation, and congruent lexicalization. This research found 28 data in total, that is, 6 insertions, 10 alternations, and 12 congruent lexicalizations. The dominant type of code-mixing used by the host is congruent lexicalization. The author also identifies the reason for code-mixing, 20 talking about a particular topic, 1 repetition for clarification, and 20 expressing group identity.

The sixth research was by Nispi and Rahayu (2023). This research focused on discovering the phenomenon of code-mixing in beauty vloggers on Instagram. This research used Muysken's theory (2000) to identify the types of code-mixing. This research used a qualitative method. This research showed that beauty vloggers use two types of code-mixing: insertion and alternation. From this research, 9 data were found, in total, 2 insertions (22,2%) and 7 alternations (77,8%) in their Instagram.

The last research was by Chen (2005). The study aims to investigate how different patterns of Cantonese-English code-mixing serve as markers of social identity among bilingual speakers in Hong Kong. This research uses the theory by Muysken (2000) typology. The studies classify intra-sentential and inter-sentential

code-mixing and examine their distinct characteristics. It examines how different age groups, education levels, and social backgrounds influence the use of specific code-mixing styles. The research highlights how speakers use code-mixing to construct and express their social identity, signaling modernity, cosmopolitanism, or educational background.

The results of this study are three important points, two distinct code-mixing styles have different social meanings, social identity and group affiliation influence code-mixing, and social identity and group affiliation influence code-mixing. Intra-sentential code-mixing (mixing English words within a Cantonese sentence) is more common among younger speakers and is perceived as a marker of modernity, global influence, and trendiness. Meanwhile, inter-sentential code-mixing (alternating entire sentences between Cantonese and English) is often used by older, highly educated individuals and is linked to higher social status and formal settings. Those with an overseas education or elite background tend to use inter-sentential code-mixing to signal higher status and international exposure.

This study also examines the different types of code-mixing used by YouTubers with education in Indonesia and Indonesia x English-speaking countries.

1.4. Research Question

Based on the background of the research, the research questions are formulated as follows:

1. What are the types of code-mixing used by the YouTubers in their YouTube channel videos?

2. Is there a relationship between the use of code-mixing and educational background?

1.5. Objectives

These researches are aimed at knowing:

1. The types of code-mixing used by the YouTubers in her videos.

2. To see the correlation between the use of code-mixing and educational background.

1.6. Scope

This study focuses on analyzing the code-mixing used by six YouTubers. There are three YouTubers with Non-English educational backgrounds: Tasya Farasya, Gel Angelica, and Shasa Zhania, and the other YouTubers who have an educational background in Indonesia x English speaking country, Caitlin Halderman, Maudy Ayunda, and Xaviera Putri. The writer chose these six YouTubers because they are famous among YouTubers and content creators, and also for the three YouTubers who have an education in Indonesia x English-speaking country, the writer chose them because the place where they study is a country that uses English with a high percentage. Indonesia x English is just a naming for the category of youtubers who have an undergraduate education abroad, this naming is given because before the college level, this youtuber also did education in Indonesia such as elementary, junior high, and also high school

Since so many videos are available on their channel, the writer limits the topic to makeup tutorials. Furthermore, the focus is only on Code-mixing from Indonesia to English. The writer only takes the first 1-2 minutes of each video to collect the data. The writer refers to Muysken (2000) in explaining the types of code-mixing. This study initially searched for code-mixing, and then classified them into multiple types. The research conducted by the writer was restricted to examining three types of code-mixing by Muysken (2000). There are (1) Insertion, (2) Alternation, (3) Congruent Lexicalization.