

# An Analysis of Code Mixing in Twitter

Pangeran Muhammad Faqih Syafaat  
*Faculty of Applied Linguistics*  
*Yogyakarta State University*  
 Yogyakarta, Indonesia  
 Pangeranmfaqih@gmail.com

Teguh Setiawan  
*Faculty of Applied Linguistics*  
*Yogyakarta State University*  
 Yogyakarta, Indonesia  
 Teguh\_setiawan@uny.ac.id

**Abstract**—The research is focused on explaining the types of code mixing that appear in twitter status. The source of the data is twitter status which find out from 8 – 24 August 2018 and the data are sentences boundary that has code mixing. The Data collection techniques used in this research is an Observation Method with writing technique. The instrument of research is human instrument and Musyken’s theory as point to classify the type of code mixing. Furthermore, the data analysis technique uses data reduction, data display and drawing conclusion. In the research finding, there was found that three types of code mixing such as insertion, alternation and congruent lexicalization. The most type which appears in twitter status is insertion and followed by congruent lexicalization and the last is alternation. The English word, phrase and clause are mixed into single sentence that has Indonesian language as a main code.

**Keywords**— *language, code mixing, twitter*

## I. INTRODUCTION

In society, the behavior and use of language are often controlled by several factors starting from existing pressure and prejudice in social environment. The language which people use in communicating everyday could represent our personality, determining the direction of our social behavior, conveying thoughts, ideas and feelings and providing more opportunities for each individual to work together and improving our development process through language [1].

It shows that a language certainly plays a very important role in human life which is not only as a medium to communicate but also have many functions. Language is a media of communication which holds very important role for human life [2]. However, many people do not pay attention to language or are often indifferent, due to the familiarity of people with language. Then, People will experience difficulties in communicating with each other without language due to the language is a tool that aims to express and deliver something from what are human mind have. Expressing anger, sadness, opinion, advice and refutation until to asking question are a few purposes of using language.

The function of language is becoming a tool for self-appreciation, communication, and conducting integration and social adaptation in certain environments or situations, and it becomes a social control [3]. There are a lot of languages that exist in the world. There are 6,909 distinct languages in the world exclude might be the language has not been documented [4]. Estimation amount of language in the world might be over 7,000 languages. It proves that

there are many varieties and diversities of language in this world.

For instance, English is an international language that have developed as second language in some Southeast Asia countries such as Singapore, Malaysia, and Philippine and also some African countries, South Africa as well as in New Zealand, Australia, Canada and the United States. In Singapore, British English is still endorsed by the government as the appropriate target variety in schools and official communications, rather than the local Singaporean English [5].

Furthermore, language is not only in oral form but also in the form of writing. Sometimes someone could not convey something clearly through oral but in writing. People could type something what is in their thought through the development of technology and information. It is increasingly advanced in the twenty first century from time to time. People could not stay away from technological advance in their daily activities. One of the technological advances which help people to communicate in writing is the use of social media. Social media has become a lifestyle nowadays. Almost everyone has a social media account to get information and communicate with each other.

The relationship between language use and social media is very close. It raises a phenomenon that mixing two languages or more in one sentence. It is also known as Code Mixing. Code mixing has become a very interesting phenomenon to be understood and studied especially in terms of the use of language in the society because it is one of essential part in the development processes and uses various languages (bilingual) or mixed (multi-lingual). In general, there are two reasons why code mixing happens in the community.

First, the speaker could only communicate and understand with the target language and second, diversity in communication [6]. Code mixing is very identical with young people because mixing language is seen as a style and brings high prestige for the users. People are encouraged to gain wider connections in the social world. Especially, youngsters in Indonesia tend to mix Bahasa Indonesia with English. Code mixing seems to be a common language practice among young people today [7].

Things like this could usually be found easily through one of the famous social media which is called by Twitter. This one of media platforms give the users full of freedom to write down what things they want to write or convey to their followers. There is no limitation in writing something both in terms of grammar and other levels except the limitation amount of character use. So that users are given

the freedom to use it. That advantage makes the use of language from twitter users are free, one of which is the use of code mixing.

The rest of this paper is organized as follow: Section II describes literature review. Section III describes material & methodology of this research. Section IV presents the research finding and following by discussion in section V. Finally, Section VI concludes this work.

## II. LITERATURE REVIEW

### A. Bilingualism

Bilingualism is the ability to use two languages even more. About half of the people in the world are native speakers who can use more than one language [8]. In line with the previous opinion, bilingualism is someone who can use 2 languages, regardless from the level of knowledge about the two languages itself [6]. However, problems come when it faced with defining the meaning of bilingualism itself because everyone has bilingual characteristics that are diverse and sometimes, they could be classified as bilingual.

Bilinguals often switch from one language to another in conversational situations that exist in everyday life. Sometimes, language switching occurs due to the circumstances and the interlocutor. For example, when someone uses a different language with a different speaker, context monitoring is needed to choose the right target language in communication and immediately switch to using the language at the right time [9]. Person who has ability to use more than one language is not only switching the language but also, they commonly mixing the language because of some reason and it could be seen the speaker of mixing is mostly teenagers.

### B. Code

In linguistics, especially sociolinguistics, the definition of "code" is different from what many common people think. Simply the term code is defined as a language. Furthermore, some scholars bring definition of code into many explanations such as definition of code is certain dialects or languages that chosen to be used on every situation, this system is usually used for communication between two or more parties, when people want to talk to each other in certain situation which has different background of language.

They must choose a certain code to express what they want to talk. The term 'code' can be used naturally to refer on any system used by two or more people to communicate each other [10]. Furthermore, this term has emerged in sociolinguistics, where it is mainly used as a neutral label for any communication system involving language - and which avoids sociolinguist must be committed to terms such as dialect, language or variation, which have special status [11].

### C. Code Mixing

Code-mixing is similar to the use or creation of pidgins, but while a pidgin is created across groups that do not share a common language. Code mixing is sociolinguistic phenomena which happen in many multilingual situation or community and it is strategy that used for convenience as well as for social interaction [12]. People have some specific

reasons why they have to mix their language such as lizard brains take over, to fit in, to get something, to say something in secret and to help conveying a thought [13].

Furthermore, another perspective sees that code mixing is often categorized as illocutionary action that is not realized by the speaker/writer in naturally occurring conversations which happen in real life and social media. However, there are no norms that guide the level of mixing of codes, different communities and people maintain it consciously or unconsciously because the level of mixing of codes is limited by the competencies possessed by each individual speaker [14].

The definition of code mixing according to experts varies greatly. Code mixing is mixing various linguistic units (morphemes, words, modifiers, phrases, clauses and sentences) especially from two grammatical systems that participate in a sentence [15]. Furthermore, code mixing appears when a person talking with the other people using two different languages but in the single utterances [10]. While another defines that Code, mixing is the mixing of a word or phrase from a language into another language in a speech community [16].

Then, the emergence of other languages in the same oral or written text, this is a general phenomenon where two or more languages are used simultaneously which is often called code mixing [14]. Another perspective reveals that code mixing is a conversation that involves intentional two different languages mixing one and another without changing related topics [17].

### D. The Types of Code Mixing

As one of interesting phenomenon in sociolinguistics, code mixing is also having several types. Even though there are some types of code mixing in many scholars offer. In this research, Muysken theory is selected because the types of code mixing are divided into three types. There are three types of code mixing such as Insertion, Alternation, and Congruent Lexicalization [18].

#### 1. Insertion

Insertion occurs when the lexical parts of a particular language insert, and it could adjust to the sentence of another language [18]. The insertion concept itself, it refers to the term used by Clyne namely transference or transfer, and the term from Myer-Scotton, which is called embedding [18]. It also appends that insertional code mixing is the constituent that is inserted; the constituent can be any syntactic unit like a lexical items or phrases. For example;

- 1) *mereka melihat **group** yang ada.*
- 2) *Dari bulan januari ke agustus saja sudah bisa **export***
- 3) *Udah lama jadi **contractor** disini, hasilnya cukup baik*

In examples 1-3 above, the words "group", then word "export" and "contractor" are English nouns inserted in the sentence. In example (1) - (3), there is one constituent of words inserted in each sentence.

2. Alternation

Alternation occurs when two languages can be replaced by one another's functions in both grammatical [18]. Alternation also occurs when two languages with different grammatical structure and lexical items alternated appearing in one sentence. The two languages appear in the clause relatively separate. It means that a clause from one language is then followed by clause another language. For example;

- 1) *Aku menggunakan tongkat for reaching these mangoes.*
- 2) *Nenek ingin kami bekerja but there is no chance for us*
- 3) *That today is my birthday ku, ingin dapet sesuatu yang special*

In examples 1-3 above, the clauses are constructed by mixing two languages. In number 1 and 2 the main clauses are Indonesia and English clauses complete the sentences while in number 3 the main clause is English and Bahasa Indonesia as a complement.

3. Congruent Lexicalization

This type refers to two languages which share grammatical structure fully or in part by shifting from one to another. It means in one sentence there are two words or phrases from each language. For example;

- 1) *Kita coba dulu sweet apa gak kalo bitter mending jangan.*
- 2) *Berbohong itu emang bad sih tapi kalo lari dari tanggung jawab crazy namanya.*
- 3) *Udah mirip cat dan dog aja pelajar sekarang jumpa langsung tawuran.*

In examples number 1 - 3, the words are more than one in each sentence. The main code that using in the sentence is Bahasa Indonesia and it is inserted by English word such as "sweet" and "bitter" in number 1, "bad" and "crazy" in number 2, and "cat" and "dog" in number 3.

III. METHODOLOGY

To conduct this research, the researcher uses descriptive-qualitative research. Qualitative research is a study to explain and to understand the phenomenon of the subject of research and the way of description the result in the form of words and sentences [19]. The source of the data is twitter status that found from 8-24 August 2018. Furthermore, the data collection technique used in this research is an Observation Method with writing technique.

The use of the method is referred because the data are obtained by reading. The term observing here is concerned not only with the use of spoken language, but also with written language [20]. The instruments of the data are human instrument and Musyken's theory as points to classify the type of code mixing. The stages of analysis will be conducted in several steps. There are data reduction, data display, and conclusion drawing/verification [21].

IV. RESEARCH FINDING

In this research, there are totally 59 data has been found. The insertion is the most found as 35 (59.32%) data. In this type of insertion, elements of English are inserted into

Indonesian sentences. The second type of code-mixing type is congruent lexicalization. In this research, the congruent lexicalization type is the second largest number data with 16 (27.12%) data. While, the findings of this research indicate that this type of alternation in code mixing is the least numbers as many as 8 (13.56%) data.

In this type of alternation code mixing, users tend to mix the clause. Indonesian clause is inserted by English clauses or vice versa in a single sentence, shown in Figure 1 and Tables I, II and III below:

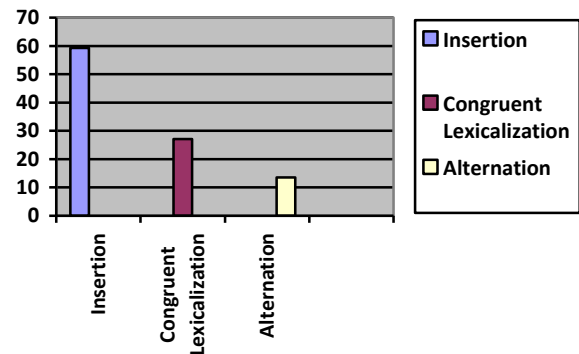


Fig. 1. The Percentage of Code-Mixing data in each type

TABLE I. THE INSERTION OF CODE MIXING DATA

No	Insertion (A)
1	<i>newbie Ini lagi mikir keras.</i>
2	<i>Lihat-lihat machine yang tanpa lelah bekerja.</i>
3	<i>Drone mahal banget harganya.</i>
4	<i>Betapa pentingnya mencabut charger ketika batrai android mencapai 100%.</i>
5	<i>Dulu dia pernah bilang di muka dosen kalau aku ini cowok yang annoying.</i>
6	<i>Kayak ada perasaan semacam kagum, heran, dan speechless gitu.</i>
7	<i>Sampai skrg gue down klo teringat omongan dia soal itu.</i>
8	<i>sosialisai SOP kaderisasi dan share kelembagaan UR.</i>
9	<i>run juga laa aku.</i>
10	<i>Ingin pergi ke Jepang but uang belum mencukupi.</i>
11	<i>Orang yg kita perdulikan blm tentu peduli dgn kita Something bad.</i>
12	<i>Positive thinking membuat seseorang lebih bahagia.</i>
13	<i>happy day ayah aku sangat sayank sama kamu semoga kau jadi ayah yang aku ingkann.</i>

TABLE II. THE ALTERNATION OF CODE MIXING DATA

No	Alternation (B)
1	<i>sepertinya acc ini akan di pakai oleh 2 orang hahaha dont worry the 2nd owner is my sista, putri.</i>
2	<i>dia mesan sesuatu when owner has been out.</i>
3	<i>Kesana kemari nyari kue but today is holiday</i>

TABLE III. THE CONGRUENT LEXICALIZATION OF CODE MIXING DATA

No	Congruent Lexicalization (C)
1	<i>Heran liat hater yg membully disosmed. Seperti mereka kurang kerjaan atau gimana ya? Ngurusin hidup artis tanpa bersudahan.</i>
2	<i>Ya kali aja dia mikir commenting sama dengan judging.</i>
3	<i>Harusnya yang namanya learner itu ya harus di teach oleh gurunya.</i>
4	<i>Slogannya sih keren" No.1 in Quality and Service.</i>

	nyatanya " <b>fake</b> ".
5	Kalau <b>money</b> ya sih bisa ngasih <b>your happiness</b> itu jadi lebih banyak dari sebelumnya.
6	<b>Good game</b> dari pemain indo semoga bisa <b>kill</b> lebih banyak lagi.

## V. DISCUSSION

### A. Insertion

Insertion data have been found on twitter as 25 words and 10 phrases. The inserted words consist of 19 nouns, 2 verbs and 3 adjectives, and 1 conjunction. The discussion will be explained as follows;

- 1) **newbie** Ini lagi mikir keras. (1A)
- 2) Lihat-lihat **machine** yang tanpa lelah bekerja. (2A)
- 3) **Drone** mahal banget harganya. (3A)
- 4) Betapa pentingnya mencabut **charger** ketika batrai android mencapai 100%. (4A)

In this insertion type, it could be analyzed that data (1)-(4) there are English words that insert into Indonesian sentences. Indonesian language is as the main code in the sentence. The word "**newbie**" is a noun in English, or it means that someone who has just started doing an activity or a job in Indonesian. Furthermore, the word "**machine**" is also English word and it means a piece of equipment with several moving parts that uses power to do a work in Indonesian language.

Then, the word of "**drone**" is an English noun which inserted into Indonesian language sentence. It means an aircraft that does not have a pilot but is controlled by someone on the ground, used especially for dropping bombs or for surveillance. Then, the word "**charger**" is a device that used to recharge a battery with electricity, and it is also an English noun. These data show that the nouns are inserted into Indonesian language sentences.

- 5) Dulu dia pernah bilang di muka dosen kalau aku ini cowok yang **annoying**. (5A)
- 6) Kayak ada perasaan semacam kagum, heran, dan **speechless** gitu. (6A)
- 7) Sampai skrg gue **down** klo teringat omongan dia soal itu. (7A)

The data (5) - (7) are still the same type of insertion because the English words are combined into Indonesian language in single sentence. However, it is slightly different because the data are adjectives in English. The word "**annoying**" is an English adjective which appears in Indonesian language sentence. In Indonesia, it could be defined as someone who slightly angry or upset. After that, the word "**speechless**" is unable to speak because someone gets angry, shocked, or surprised. It is also English adjective. And the last is the word "**down**". It is English adjective and it means that feeling which sad because of something.

- 8) sosialisai SOP kaderisasi dan **share** kelembagaan UR. (8A)
- 9) **run** juga laa aku. (9A)

The data showed that the word "**share**" and "**run**" are could be categorized as English verb. These verbs are

inserted into single Indonesian sentences. Data number (8) is share which has a meaning as if two or more people share an activity, they each do some of it in Indonesian language while the word "**run**" is actually to move along, faster than walking, by taking quick steps in which each foot is lifted before the next foot touches the ground.

- 10) Ingin pergi ke Jepang **but** uang belum mencukupi. (10A)

Insertion English conjunction to bridge Indonesian clause could be seen in data above. The data (10) the conjunction "**but**" is known as coordinate conjunction in English. There is mixing between English conjunction and Indonesian sentence

- 11) Orang yg kita perdulikan blm tentu peduli dgn kita **Something bad**. (11A)
- 12) **Positive thinking** membuat seseorang lebih bahagia. (12A)
- 13) **happy day** ayah aku sangat sayank sama kamu semoga kau jadi ayah yang aku inginkann. (13A)

Insertion is not only in word level but also in the level of phrases. The data (11) has English phrase in Indonesia language sentence. It could be seen in the phrase "**something bad**" which inserts into the sentence. It could be defined as something that a person does not expect in Indonesia. Furthermore, data (12) also has an English phrase. That is "**positive thinking**". It means that a way of thinking that produces good thoughts and attitudes and makes people enthusiastic. Furthermore, data (13) the phrase is "**happy day**" and it is definitely English phrase which has meaning as the day that bring joy in Indonesia language.

### B. Alternation

Alternation occurs when there are mixing codes between clause boundaries. Alternation data have been found on twitter as 8 clauses. The discussion will be explained as follows;

- 14) *sepertinya acc ini akan di pakai oleh 2 orang hahaha* **dont worry the 2nd owner is my sista, putri**. (1B)
- 15) *dia mesan sesuatu* **when owner has been out**. (2B)
- 16) *Kesana kemari nyari kue* **but today is holiday** (3B)

In the data above, data (14) there is clause in Indonesia and English clause. Furthermore, the data (15), it could be showed that there is independent clause in Indonesia language such as "*dia memesan sesuatu*" and it is followed by English clause such as "*when owner has been out*" which is dependent clause. Then, data (16) is also showing that there are two clauses in different languages such Indonesian and English clause. "*Kesana kemari nyari kue*" is Indonesian clause while "*but today is holiday*" is English clause which stands as dependent clause. In this type, there is clearly boundary which make this type appears into code mixing.

### C. Congruent Lexicalization

Congruent lexicalization data have been found on twitter as 16 data. The discussion will be explained as follows;

- 17) Heran liat **hater** yg **membully** disosmed. Seperti mereka kurang kerjaan atau gimana ya? Ngurusin hidup artis tanpa bersudahan. (1C)
- 18) Ya kali aja dia mikir **commenting** sama dengan **judging**. (2C)
- 19) **Harusnya** yang namanya **learner** itu ya harus di **teach** oleh gurunya. (3C)

Congruent lexicalization is the type of code mixing which occurs when there is more than one word or phrase in a single sentence. In this case, two English words are inserted into each Indonesian language sentences. In data (17) there are two English words such as “*hater*” and “*membully*”. The word “*hater*” is a single constituent which refers to an English noun. It means someone who does not like something or someone else in Indonesia. It is embedded in single sentence that having Indonesian language as the main code. However, another word “*membully*” is actually combination or hybrid between Indonesian prefix “*mem-*” and English word “*bully*” which refers to English verb.

The word “*bully*” is using superior strength or influence to intimidate someone. Furthermore, in data (18) “*commenting*” and “*judging*” are showed as two English words that inserted into Indonesian sentence. These are English verb. Commenting is known as the way someone to express opinion or something which aimed to someone or something else, it could be positive or negative expression while judging is form of opinion or conclusion about. The last is data (19) “*learner*” and “*teach*” are also English word. These two English words are made part of a sentence. “*Learner*” is someone who learns about something while the word “*teach*” is showing or explaining to someone how to do something. From analysis above, it could be seen the pattern of congruent lexicalization lies in two English words which put into single sentence. The word boundary is clearly seen in this type.

- 20) **Slogannya** sih keren” No.1 in **Quality and Service**, nyatanya **fake**”. (4C)
- 21) **Kalau money** ya sih bisa ngasih **your happiness** itu jadi lebih banyak dari sebelumnya. (5C)
- 22) **Good game** dari pemain indo semoga bisa **kill** lebih banyak lagi. (6C)

The data (20) showed that there are single phrase and word in a sentence. The English phrase is “*in Quality and Service*” while the English word is “*fake*” which refers to adjective. Furthermore, data (21) also has one word and phrase. There is “*money*” that also as a noun in English and “*your happiness*”. These cases could be seen clearly pattern that congruent lexicalization is not always put only English word into the Indonesian sentence but also mix one English phrase and word into single Indonesian sentence. Then, in data (22) has English phrase and word in Indonesian sentence. There are “*Good game*” and “*Kill*” which refers to verb. This is different from analysis above before because in single sentence has English phrase and word. English phrase and word are mixed into the Indonesian language sentence.

## VI. CONCLUSION

Types of code mixing that has been found are Insertion, Alternation, and Congruent Lexicalization. In this research, it showed that the most type of code mixing is insertion while congruent lexicalization is in the second and the last is alternation. Insertion could be seen that the English words are inserted by users of twitter into single sentence that dominated by Indonesia language. It was not only in word boundary but also in Phrase. Phrase is still inserted into Indonesian language sentence as main code which uses by users. In this type, it could be seen if insertion is always having limitation in inserting one word or phrase in a sentence. Furthermore, congruent lexicalization showed that word or phrase that appears or mix into single sentence is more than one word or phrase. It is different from insertion that only having one word or phrase in sentence. Either English word or phrase has been inserted in more than single both of it. It is possible that there are two English words in single sentence or two phrases even one-word side by side with phrase in a sentence which Indonesian language as main code. Congruent lexicalization inserted word or phrase more than single level in a sentence. Then, alternation appears in clause boundary between different languages in Indonesian and English. The English clause which is used has different position. It could be Independent clause or dependent clause in a sentence.

## REFERENCES

- [1] S. Chaturvedi, "A Sociolinguistic Study of Linguistic Variation and Code Matrix in Kanpur", *Procedia - Social and Behavioral Sciences* 192, 107 – 115, 2015.
- [2] Anderson, S.R, *Languages: A Very Short Introduction*, Oxford University Press, Oxford, 2012.
- [3] Ayeomoni, M.O. "Code-Switching and Code-Mixing: Style of Language Use in Childhood in Yoruba Speech Community", *Nordic Journal of African Studies*, 15(1), 90–99. 2006.
- [4] Gorys, K, *Komposisi Sebuah Pengantar Kemahiran Bahasa*, Nusa Indah, Jakarta, 1994.
- [5] Gumperz, J.J, "Studying language, culture, and society: Sociolinguistics or linguistic anthropology?", *Journal of Sociolinguistics*, 12(4), 532–545, 2008.
- [6] Sumarsih, M.A., S. Bahri, & D. Sanjaya, "Code Switching and Code Mixing in Indonesia: Study in Sociolinguistics", *English Language and Literature Studies*, 4(1), 77-92, 2014.
- [7] B. Kurniawan. "Code-mixing on Facebook postings by EFL students: A small scale study at an SMP in Tangerang". *Indonesian Journal of English Language Teaching*, 11, 2, 169-180, 2016.
- [8] Fromkin, V. *An Introduction to Language*. Michigan University. 2003.
- [9] A. d. Bruina, A. G. Samuela, J. A. Dunabeitiad, "Voluntary language switching: When and why do bilinguals switch between their languages?", *Journal of Memory and Language*. vol. 103, 28–43. 2018.
- [10] Holmes, J, *An Introduction to Sociolinguistics*, Longman Publishing, New York, 1992.
- [11] Crystal, D, *A Dictionary of Linguistics and Phonetics Sixth Edition*, Blackwell Publishing, USA, 2008.
- [12] T.K. Bhatia, & W.C. Ritchie, *The Handbook of Bilingualism*, Wiley, 2004.
- [13] Mujiono, R. Wilujeng, M. Suharto, "Code Mixing as a Communication Strategy Performed by Outbound Call (OBC) Center

- Agents”, *International Journal of Social Sciences & Educational Studies*, 3(3), 4-18, 2017.
- [14] Wardhaugh, R. *An introduction to sociolinguistics*. Cambridge MA: Blackwell, Cambridge, 1998.
- [15] Muysken, *Bilingual Speech: A Typology of Code Mixing*. Cambridge University Press, Cambridge, 2000.
- [16] N. Yuliana, A.R Luziana, & P. Sarwendah, “Code-Mixing and Code-Switching of Indonesian Celebrities: A Comparative Study”, *Jurnal Lingua Cultura*, 9(1), 47-54, 2015.
- [17] A. M. ul-Alam, & S. Quyyum, "A Sociolinguistic Survey on Code Switching & Code Mixing by the Native Speakers of Bangladesh". *Journal of Manarat International University*, 6(1), 2016.
- [18] Muysken, *Bilingual Speech: A Typology of Code Mixing*. Cambridge University Press, Cambridge, 2000.
- [19] Moleong, L.J. *Metodologi Penelitian Kualitatif*. Remaja Rosda Karya, Bandung, 2010.
- [20] Alasuutari, P. (1995). *Researching culture: Qualitative method and cultural studies*. Sage.
- [21] Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif dan R&D*, Alfabeta, Bandung, 2009.