Social Aspects in Social Media: Code Switching and Code Mixing

In Twitter

Alaa Hussein Samhan

Department of Arts, Zarqa University, PO box 2000 - Zarqa 13110 Jordan

* E-mail of the corresponding author:alaasamhan87@yahoo.com

The research is financed by the Deanship of Research in Zarqa University /Jordan

Abstract

This study investigated if there was code mixing or code switching in one hundred tweets (from April 8th to May 15th) of eighty participants had an active accounts on twitter and if age, gender, and education had any role in affecting the participants language. In this research the researcher divided the participants into three groups regarding to their age, gender and educational level. Then analyzed their tweets referring to the existence of code mixing or code switching. According to that the study revealed that age and education could firmly affect the participant language while gender couldn’t.

Key words: code mixing, code switching, twitter, tweet.

1. Introduction

So many studies have studied how age, gender and education could affect language from different perspectives such as loan words, language acquisition, and bilingual students language, etc. This study will investigates if these three aspects (age, gender and education) could affect the use of code mixing and code switching in language used by the participants in twitter. The most important thing that this study built on sociolinguistics which defined according to fishman the study of language varieties characteristics, their speakers and their functions within a speech community. While Holmes argues that it is how people use language in different contexts and how to indicate their social identity by language and he says that it is a developed science from the cooperation between sociology and linguistics which investigates the language social meaning, Florian Coulmas (2008).

Code mixing and code switching are branches of sociolinguistics where code mixing according to Crdenas carlos (2009) happened when a person spoke two or more languages and it is impossible when the speaker is monolingual. While Code switching As Gal (1988, p. 247) says, ‘is a conversational strategy used to establish, cross or destroy group boundaries; to create, evoke or change interpersonal relations with their rights and obligations’. Code-switching occurs in conditions of change, where group boundaries are diffuse, norms and standards of evaluation vary, and where speakers’, Ronald Wardhaugh (2014).

This study aims to investigate if age, gender, and education could affect language code switching and code mixing used in twitter.

2. Questions

1- Does age affect the language used in twitter?
2- Does gender affect the language used in twitter?
3- Does education affect the language used in twitter?

3. Literature review

3.1 Previous studies

Age, gender and education affect language used in a certain society. Nowadays they also affect language used in social media which become a wide domain for using language. Many recent studies have studied these social factors and how they affect language by studying the linguistic features of the language used in a certain domain.

Cristan Mizil, Michal Gamon and Susan Dumais (2011) argued in their study the psycholinguistic theory of communication accommodation accounts in the context of twitter conversation. They developed a probabilistic frame work that can model accommodation and measure its effects then applying it to a large twitter conversational data set. the researchers discover a complexity of the phenomenon and they also
observed the potential relation between stylistic influence and network features commonly associated with social status.

Azadeh Nemati and Jennifer Marie Bayer (2007) discussed whether men and women language were different with respect to the use of intensifiers, hedges and tag questions in English and Persian. The researchers depend on Robin Lakoff’s study which results that gender differences in language usage reflect different and unequal roles and status. The results obtained in this study indicate that Lakoff’s ideas concerning tag questions, hedges and intensifiers cannot be upheld, given the corpus under study and the three null hypotheses stated earlier can be upheld.

H. Andrew Schwartz and 10 other researchers (2013) shed light on psychosocial processes yielding results that are face valid, tie in with other research, suggest new hypotheses, and give detailed insights. The researchers analyzed 700 million words, phrases, and topic instances collected from the Facebook messages of 75,000 volunteers, who also took standard personality tests. This study found striking variations in language with personality, gender, and age.

Emad Al-Den Omar (2013) where he aimed in his study to identify the linguistic features of Arabic language in social media and to determine whether the language used there was standard or not. The researcher used a Descriptive analytical method where he analyzed the language used in messages within different social media. The study found that the language used in social media somehow related to the standard language in some vocabularies and these vocabularies differ from one person to another according to education and occasion.

Neny Isharyanti & Monica Stella Crdenas-Claros (2009) where they examined code mixing and code switching In a chat room conversations of 12 non-native speakers of English from Spanish and Indonesian backgrounds were collected during a two month period and analyzed. The study found that technology-related terms, along with introductory terms, triggered more instances of code switching and code mixing regardless of the linguistic background of the participants.

Ronald Wardhaugh (2014) a book intended to provide students with a sound, basic coverage of most of the topics dealt with in courses described as either ‘Sociolinguistics’ or ‘The Sociology of Language.’ It assumes very little previous knowledge of linguistics, anthropology, or sociology.

Florian Coulmas (2008) his book content is considering the fundamental topics of its two main sections; Part 1: Micro-choices (from Chapters 2-6) and Part: 2 Macro-choices (Chapters 7-13). As the title itself suggests the book predominantly focuses on the notion of choice and its effects on communicative speech.

3.2 Theoretical back ground

3.2.1 Sociolinguistics

Florian Coulmas (2008) mentioned in his book defining sociolinguistics that it is the study of language varieties characteristics, their speakers and their functions within a speech community. According to Holmes he argued that it is how people use language in different contexts and how to indicate their social identity by language. Also, he says that it is a developed science from the cooperation between sociology and linguistics which investigates the language social meaning.

3.2.2 Code mixing

Code mixing happens when a person speaks two or more languages and it is impossible when the speaker is monolingual. It also called intra sentential code switching according to Crdenas carlos (2009). There are three types of code mixing according to Muysken (2000) which are insertion, alternation, and congruent lexicalization.

3.2.3 Code-switching

As Gal (1988, p. 247) says, ‘code switching is a conversational strategy used to establish, cross or destroy group boundaries; to create, evoke or change interpersonal relations with their rights and obligations’. Code-switching occurs in conditions of change, where group boundaries are diffuse, norms and standards of evaluation vary, and where speakers’ regarding to Ronald Wardhaugh (2014).

Chad Nielp (2006) mentioned according to poplack (1980:605) he defines three types of code switching:

1- Tag switching which defined according to Romaine (1995): “subject to minimal syntax restriction” where tags reflect lower language proficiency in comparison with inter and intra switching.

2- Inter-sentential switching it is according to (Myers-Scotton 1993:3) the switching between sentences from one language to another which produced entirely in one language before switching to another.

3- Intra-sentential switching occurs within the same sentence or a part of it (Myers – scotton 1993:4). According to Gumperz (1982) he classified the code switching into:

1-Situational: when the speaker uses two different codes for two different situations. That means the control factor here is situation.

2-Metaphorical: when the speaker chooses the language according to the topic. That means the control factor here is topic.
### 3.2.4 Twitter and tweet

According to Alex Wilson, Phil Blunsom, and Andrew D. Ker (2014) “Twitter is a social networking platform, launched in 2006. Users of the service post short messages of up to 140 characters (tweets); this type of messaging is known as microblogging. By default, these tweets are publicly visible: it is not even necessary to have a Twitter account to read them. Each user has a username, and tweets can be directed at another user by placing ‘@’ in front of their username somewhere in the tweet. A unique feature to Twitter is the ability to re-tweet a message; re-posting another user’s message to spread the message, echo the sentiment, or add a comment. Such tweets are usually marked as such with ‘RT’ placed at the start of the tweet.”

Tweet is a message sent on Twitter. To send or receive a Tweet, you have to create a free account with Twitter. You also need to have friends and contacts with Twitter accounts; otherwise you’re typing to the void. (http://computer.howstuffworks.com/internet/social-networking/networks/twitter1.htm)

### 4. Methodology

The researcher used the analytical approach to collect a random sample from twitter which was eighty participants who had active profiles. Age, gender and education as data of twitter users were collected from their profiles information in twitter. Otherwise, when there was a missing data about these three aspects, the researcher used their Facebook profiles if it was mentioned in their Twitter profiles, asked them in their Ask profiles if they had one or messaged them in Facebook or Twitter.

Each group was divided into two categories or more. Gender group was divided into two categories, males and females in order to analyze the data then compared with each other referring to the existence of code switching and code mixing. As well as education group were classified under two categories, high educational level and low educational level, where the researcher investigated if their tweets presented code mixing or code switching then compared the analyzed data to revealed results. On the other hand, age group had three age categories which involved participants ages which were (15-25), (25-35) and (over 35) years old that also analyzed and compared as the previous two groups.

One hundred tweets of them were taken from the period of (April, 8th to may, 15th) using a mobile phone and screen shoot application. After that they were divided into three groups according to age, gender, and education. Then the researcher analyzed the data (tweets) according to the existence of code mixing and code switching. After that the researcher compared the analyzed data of each group categories with each other regarding to the code mixing and code switching presented in their tweets.

### 5. Analysis and Discussion

This tweet of the participant hadn’t any code switching or code mixing because she used one variety of a language which were colloquial Arabic.

Here the tweet also hadn’t any code switching or code mixing where he used one variety of a language which were standard Arabic.

While here the participant shown code switching because the participant used two different languages, where he switched from colloquial Arabic (a variety of Arabic language) to English language.

Best show begins in five minutes @ ro’ya tv #

But it was noticed in this tweet that it had code mixing because the participant used two different languages at the same time, English and Arabic, where she started with English and ended with Arabic.

If we had a look to this tweet, we noticed that the participant used two varieties of one language (Arabic) where he switched from standard to Jordanian colloquial Arabic.

The researcher found that the tweet of the participant had code mixing because he used two different varieties of Arabic language (standard and colloquial) and one different language (English language) at the same time.

My brother is such a weirdo, oh my god I love him hahaha

Comparing with the previous tweet, it was noticed that there is no existence of code mixing or code switching because the participant used one single language which was English.
On the other hand, this tweet had code switching between two dialects of one language (Arabic) which was Jordanian colloquial Arabic and Egyptian colloquial Arabic, where the participant started by Jordanian dialect and ended with Egyptian dialect.

As well as, this tweet had code switching between two languages (Arabic and English) which was Jordanian colloquial Arabic and English language, where the participant started by Jordanian dialect and ended with English term.

I don’t think sara7a ma fe mosta7el

This tweet had the same function of the previous one but they differed in the languages which were switched from and switched to, where the participant started by English and switched to Jordanian Arabic.

Comparing with the previous tweet, this tweet had the same function where the participant switched from English language to Arabic language.

The tweets of the participants hadn’t any code switching or code mixing because the participant used one language which was standard Arabic language without any other dialects or languages.

On the other hand, the researcher found in this tweet code switching between two languages (English and Arabic), where the participant started standard Arabic then switched to English and returned to standard Arabic.

6. Results

After the researcher collected, analyzed and studied the data, findings were revealed and scheduled as followed to answer the research questions.

6.1 Does age affect the language used in twitter?

To answer this question the researcher took the age group and after analyzing the tweets of its categories, she compared the analyzed data with each other referring to data which had code switching and other data which had code mixing. Then, the researcher found that participants of the age period (15 – 25) had code mixing and code switching in their tweets largely but code mixing was used more than code switching. These findings didn’t show that this group knew more than one language or had a high educational level. It may indicate that they mixed and switched for fashion or stylish purposes. Regarding to the age period (25-35), the researcher noticed that code mixing and code switching were few existed but code switching was used more than code mixing. Although most of them were high educated and knew more than one language, they almost use them in their tweets. These findings may indicate that they mixed or switched for argent things such as foreign names, places and scientific terms or to clarify things. Also, the researcher found that in age period (over 35) code mixing and code switching rarely existed and the participants didn’t tend to use them. Schedule (1) elucidated the findings that mentioned above.

<table>
<thead>
<tr>
<th>Age</th>
<th>Result</th>
<th>28,57% switched</th>
<th>57,14% mixed</th>
<th>14,28% didn’t mix or switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>(15-25)</td>
<td>code mixing and code switching were existed largely but code mixing was used more than code switching.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(25-35)</td>
<td>Code mixing and code switching were few existed but code switching was used more than code mixing.</td>
<td>40,62% switched</td>
<td>15,62% mixed</td>
<td>43,75% didn’t mix or switch</td>
</tr>
<tr>
<td>(over 35)</td>
<td>code mixing and code switching were rarely existed and the participants didn’t tend to use them.</td>
<td>14,28 % switched</td>
<td>6,25 % mixed</td>
<td>78,12% didn’t mix or switch</td>
</tr>
</tbody>
</table>
6.2 Does gender affect the language used in twitter?

The researcher divided this group into two categories, males and females. In order to answer this question, categories data were analyzed then compared with each other referring to the existence of code switching and code mixing. Then, the researcher found that males tended to mix and switch in their speech because they had to do it according to their topics such as sports and cars. While females also tended to mix and switch in their tweets because they think that this is more prestigious and stylish. Then, gender didn’t affect the use of code mixing and code switching as the researcher thought, because both males and females switched and mixed in their tweets but the aim was differed. Schedule (2) elucidated the findings that mentioned above.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Result</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Males tended to mix and switch in their speech because they had to do it according to their topics such as sports and cars.</td>
<td>60% switched and mixed</td>
</tr>
<tr>
<td>Female</td>
<td>Females also tended to mix and switch in their tweets because they think that this is more prestigious and stylish.</td>
<td>40% switched and mixed</td>
</tr>
</tbody>
</table>

6.3 Does education affect the language used in twitter?

Yes, it does. The researcher achieved to this answer after studying and analyzing the data regarding to classifying it under two categories, high educational level and low educational level. High educational level participants mix and switch rarely although most of them knew two languages or more. The researcher thought that maybe because they more aware about the fact of language regression. On the other hand, Although the low educational level participants group were low level educated and knew just one or two languages but the researcher found that they switched and mixed more than high educated people. This maybe because they less aware about the fact of language regression. Schedule (3) represented the results mentioned above.

<table>
<thead>
<tr>
<th>Education level</th>
<th>Result</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>High level</td>
<td>High level educated participants mixed and switched rarely although most of them knew two languages or more.</td>
<td>11.2% switched and mixed</td>
</tr>
<tr>
<td>Low level</td>
<td>Although this group of participants was low level educated but they switched and mixed more than high educated people.</td>
<td>88.3% switched and mixed</td>
</tr>
</tbody>
</table>

7. Conclusion

After the researcher collected, analyzed and studied the data, the study revealed the following:
1-Age could firmly affect the participant language.
2-Gender didn’t affect the participant language.
3-Education affected the participants language.

8. Recommendations

After studying language used in the social media website "twitter” the researcher recommended that:
1- Youth and low educated people must be aware of the bad impact of code switching and code mixing on language. Code switching and code mixing are not a measure of our education.
2- Universities must have a role in making youth students more aware of using the language in the right way and to be proud of their language in order to not reduce the language outcome.

3- Social media and multimedia must had awareness role by presenting this issue and reducing it.

4- For future studies it may preferable to study the bad effect of code mixing and code switching on language especially language used by youth people who will carry the language for the next generation.

References
- Emad Al-Den Omar, Arabic language and social media, the 2nd international conference of Arabic language, Beirut, (2013).
- H. Andrew Schwartz and 10 other researchers, Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach, University of Warwick, United kingdom, (2013).
- www.ask.com
- www.facebook.com
- www.twitter.com