6-1-2016

Code-choice on Twitter: How stance-taking and linguistic accommodation reflect the identity of polyglossic Egyptian users

Sahar Mashhour

The American University in Cairo

School of Humanities and Social Sciences

Code-Choice on Twitter: How Stance-taking and Linguistic Accommodation Reflect the Identity of Polyglossic Egyptian Users

A Thesis Submitted to

The Department of Applied Linguistics

In partial Fulfillment of the Requirements for

The degree of Master of Arts

By

Sahar Mashhour

Under the supervision of Dr. Reem Bassiouney

Dr. Amira Agameya

Dr. Atta Gebril

May 2016

The American University in Cairo

School of Humanities and Social Sciences (HUSS)

Code-Choice on Twitter: How Stance-taking and Linguistic Accommodation Reflect the Identity of Polyglossic Egyptian Users

A Thesis Submitted by

Sahar Mashhour

Submitted to the Department of Applied Linguistics

May 2016

In partial fulfillment of the requirements for
The degree of Master of Arts
in Teaching English to Speakers of Other Languages TESOL

has been approved by

Dr. Reem Bassiouney _____
Thesis Supervisor
Affiliation: Department of Applied Linguistics
Date May 11, 2016

Dr. Amira Agameya _____
Thesis first Reader
Affiliation: Department of Applied Linguistics
Date May 15, 2016

Dr. Atta Gebril _____
Thesis Second Reader
Affiliation: Department of Applied Linguistics
Date May 11, 2016

Dr. Amira Agameya _____
Department Chair
Date May 15, 2016

Dr. Nathaniel Bowditch _____
Dean of HUSS
Date 24/5/2016

ACKNOWLEDGEMENTS

I would like to express my gratitude to my very dear professors who have given me the support and guidance I needed during these two years of my life. Dr. Amira Agameya, I shall forever remain grateful for believing in me, for being such an amazing academic advisor, and for encouraging me to pursue my MA degree at AUC. You are a true role model I am looking up to. Thank you, Dr. Reem Bassiouney, for your time and help. This study would not have been what it is without your input and your time which you never hesitated to offer me. Dr Atta Gebril, you were my first, my second, and my third reader at different points over these two years. Thank you for being a great example of a dedicated professor whose office was always open to all the students who sought his advice or help.

I am lucky to have wonderful friends like Nourhan Sorour, Rasha Soliman, and HebatAllah Mohamed who have never hesitated to help me and were always there for me. Nesma Hossam and Noura Nabil, it would have been much harder if I had gone through this journey alone. I am very grateful that you were with me and by my side all along. Tina Francis, thank you for your help with the translations needed in the study and for being the kind of person and friend you are. Alyssa Young, I am indebted to you for copy-editing this study. You are an angel!

Finally, my caring and loving family, it was a great comfort knowing that at the end of any rough day, I have your unconditional love and encouragement which kept me going and pushed me beyond my limits. Hadeel Mashhour and Omar Mashhour, thank you for your help with Twitter and the choice of tweets.

This study is dedicated to you, mommy. I love you so much!

ABSTRACT

This study examines the online identity of polyglossic Egyptian users of Twitter. It is descriptive and exploratory utilizing a qualitative design with some frequency count which adds descriptive data. Data were collected using a Discourse Completion Task (DCT) where the participants were presented with a number of tweets and were asked to type another tweet in response to each. The findings from the study suggest that polyglossic Egyptians, those who are proficient in English as well as Arabic, exhibited an assertive identity on Twitter. This identity was constructed through the choice of code, the linguistic accommodation to the tweet authors, and the stance they took. Polyglossic Egyptians were found to use English more than any other code, followed by Arabizi, and then Arabic. They linguistically accommodated the tweet authors in their replies to some extent by choosing the same code in replying as that used in the original tweet. Further, and using Du Bois' (2007) stance triangle framework, it was also found that they expressed their (dis)alignment quite bluntly by taking an epistemic stance achieved through the use of boosters (very few hedges were used), sarcasm, simple present tense (to express an opinion as if stating a fact), and modals (to offer advice). By doing that, polyglossic Egyptians were found to be assertive in expressing their opinions, often showing themselves as informative, superior people who are guided by facts about topics rather than feelings.

Keywords: accommodation, Arabizi, diglossia, discourse 2.0, identity, micro-blog, polyglossia, stance, Twitter

TABLE OF CONTENTS

LIST OF TABLES

TRANSCRIPTION

The following table shows the transcription symbols used by Bassiouney (2010). The same symbols are used in this study.

Table 0.1

The Pronunciation of the Letters of the Arabic Alphabet

ا	*ʔ / a / a:*	ذ	*Đ*	ظ	*ẓ*	ن	*n*
ب	*B*	ر	*R*	ع	*ʕ*	ه	*h*
ت	*T*	ز	*Z*	غ	*Ġ*	و	*w / u / u:*
ث	*Θ*	س	*S*	ف	*F*	ي	*y / i / i:*
ج	*J / g*	ش	*ʃ*	ق	*Q*		
ح	*ḥ*	ص	*ṣ*	ك	*K*	ة-	*-a*
خ	*X*	ض	*ḍ*	ل	*L*		
د	*D*	ط	*ṭ*	م	*M*	ء	*ʔ*

LIST OF ABBREVIATIONS

AE	Arabized English
DCT	Discourse Completion Task
CMC	Computer-mediated Communication
CMD	Computer-mediated Discourse
CMDA	Computer-mediated Discourse Analysis
CS	Code-switching
DA	Discourse Analysis
DM	Direct Message
EA	Electronic Amiyya
ECA	Egyptian Colloquial Arabic
MSA	Modern Standard Arabic
SA	Standard Arabic
SEE	Sounds, Emoticons, and/or Emojis

CHAPTER ONE: INTRODUCTION

Sociolinguistics as a branch of the larger field of Applied Linguistics is concerned with the study of language in use within context. Sociolinguists perceive society and people as the driving force that set Sociolinguistic research into action. With the advent of the internet and the late twentieth century technology, people have access to a different world that had never existed before. This world has a life of its own, a life where communication, language, media, and awareness carry different meanings. This world is called Web 2.0, and it is defined as web-based platforms which generate interaction by users (Herring, 2011). This type of interaction has resulted in Discourse 2.0—that is integral to our communication and which takes place along with present interaction.

Discourse 2.0 is emerging as a subject of study in the field of sociolinguistics and will be the focus of the proposed study. One of the terms key to this new discourse is *social media*. Page, Barton, Unger, and Zappavigna (2014) define social media as a term that refers to internet websites which are built upon the interaction between participants and whose content is published by people on that medium and is seen by the audience these people choose. Bassiouney (2014) notes that the difference between old forms of media and new online media resides in the fact that online media allow people to interact and express their opinions rather than merely being receptive of others' opinions. Bassiouney (2010) adds that for someone to create an identity, one has to express himself or herself. Hence, language whether written or spoken is a significant indicator of a person's identity, since it is the means by which people projects themselves (Bassiouney, 2009).

One way of explaining the relationship between one's use of language and one's identity is stance. Stance is defined by Bassiouney (2014) as how one positions himself or herself in

relation to others, situations, or things. This positioning could be mediated through the code people use. Du Bois (2007) explains stance through what he calls a *stance triangle* comprised of three components: evaluation, positioning, and alignment. By taking a stance, people evaluate something and thereby position themselves with regards to it and either align or dislaign themselves with other(s) (Du Bois, 2007).

Social networking platforms are one of the factors which have resulted in the spread of a code of writing which has unique characteristics. This code is called Arabizi. It is writing Arabic in Latin letters (Yaghan, 2008). Many studies explored the development of this code (Abu Elhij'a 2012; Muhammed 2013; Yaghan 2008), and they found that there were different reasons for its development including faster typing, the ability to add English words without having to switch the keypad, and a better communication of one's feelings through capital letters for example.

Twitter is one of the widely used social media platforms. It is considered, and sometimes thought of as synonymous to, a micro-blog in the sense that its users post their updates (which are referred to as *tweets*) with a limit of only 140 characters. Thus, Twitter is an interesting area to study language: if one has such a limited number of characters for self-expression, then examining how people express themselves and how their linguistic choice reflects their identities, especially if they have access to more than one code, would be indeed interesting to explore.

In a closely-related study, Kosoff (2014) examined code-switching on Twitter in Egypt. In her study, Kosoff described Egypt as a polyglossic environment, one where varieties of Arabic and foreign languages are used. When a community speaks more than one variety of a language along with another language, it is called a polyglossic community. The codes Kosoff found to be used were English, Modern Standard Arabic (MSA), Egyptian Colloquial Arabic (ECA), Arabizi

MSA, Arabizi ECA, and Code-switching (CS) between any two of these codes. Kosoff examined the tweets of 10 Twitter users over a period of time and counted the frequency of use of each of these codes in their tweets. She then explained that their code-choice depended on the kind of audience they were targeting.

Statement of the Research Problem

The literature available discussing identity in relation to social media is not vast. What is available is either focused more on blogs or political topics, triggered by the rise of political activism which the world is witnessing (Elsadda, 2010) or on the functions of technicalities like the *hashtag* and the *retweet* on Twitter. Elsadda (2010) highlighted that the kind of influence of cultural blogs on aspects of life that were not political was yet unclear since very little research has been done in that area. More research is therefore needed on micro-blogs, particularly research with a social focus on people and how they express themselves.

Stance has been examined extensively in academic writing (Charles 2006; Hyland 2005; Silver 2003), in conversational discourse (Bassiouney 2015; Jaffe 2009; Johnstone 2007; McCann, Ota, Giles, & Caraker 2009) and in corpora of both (Englbretson 2007; Hunston 2007; Rauniomaa 2007). However, it has not been studied in Discourse 2.0, nor on Twitter specifically.

Further, polyglossic Egyptian users of Twitter are under-represented in the literature except for Kossoff (2014), the only study which explored the choice of code by polyglossic Egyptians. The scarcity of research in this area is one of the reasons for undertaking this study.

Purpose of the Study

What this study aimed at investigating was how polyglossic, Egyptian users of Twitter used different codes to express their ideas as they replied to tweets. Examining their code choice

and the stance they took was used to understand their online identity and how it is shaped by these factors.

Research Questions

The current study aimed at looking into the codes and the methods through which polyglossic Egyptians take stance on Twitter through their tweet replies. The research questions answered in this study are:

1) What are the codes used by polyglossic Egyptians in their replies to tweets on Twitter?
    a. How frequently do they use each of these codes?
    b. How is the choice of code related to the stance taken on Twitter?
    c. What is the relation between code-choice and the religious and political identification?
2) What is the relation between stance and accommodation on Twitter?
3) What are the replying and the stance-taking strategies exhibited by polyglossic Egyptians on Twitter?

Delimitations

In an attempt to fill in the gap in the literature, codes used by polyglossic Egyptians, those who are proficient in English as well as Arabic, in their tweet replies were examined. The study did not aim at investigating the political context. Rather, it attempted to understand the relationship between the code choice and the users' identity through examining the stance they took.

Definition of Constructs

Theoretical Definitions.

Accommodation: Shifting our style according to the audience and/or according to the topic, either consciously or unconsciously. Accommodation, according to Johnstone (2008), occurs through code choice, accents, and also though "stylistic and rhetorical decisions on all levels" (p. 148).

Affective stance: Du Bois (2007) explains that people take an affective stance when they "choose a position along an affective scale" (p. 143). In other words, it occurs when the stancetaker shows his or her emotions about the object of stance. Stance markers for an affective stance are adjectives like *glad* and *amazed*.

Alignment: It is the act of establishing a relationship between two stances taken by stancetakers through agreement or disagreement (Du Bois, 2007, p. 144). Stancetakers show alignment through markers, like *yes*, *no*, and *too*, and through verbs like *agree*.

Arabizi: Yaghan (2008) defines it as an Arabic language system which is written using Latin characters and which is formed from the combination of the two Arabic words "'arabi' (Arabic) and 'engliszi' (English)." It has been developed in the Arab world following the technological advances the world is witnessing and acts as the means by which people are able to communicate freely using their own codes (in Egypt, ECA) without having to switch to a more perceived formal form of writing (in Egypt, MSA) which they do not use in their everyday communications (Abu Elhij'a, 2012).

Boosters: Devices like *certainly* and *obviously* which reflect how positive a writer is about the information presented (Hyland, 2005). Hedges and boosters help writers determine where they would like their words to fall on a scale ranging from fact to opinion. Boosters

shut the space for dispute and invite readers instead to a more reliance on the writer's words (Hyland, 2005).

Code: It is a neutral term many linguists, as well as the researcher in the current study, prefer to use when they refer to a certain language, dialect, or variety instead of using other biased terms that imply a judgment on the status of that code.

Code-switching (CS): It is defined by Myers-Scotton (2010) as the shift in codes people who have access to more than one language make during their speech, but it is not only limited to speech. It takes place in writing as well.

Computer-mediated communication (CMC): It is any form of communication that is based on computer as a medium for it.

Computer-mediated discourse (CMD): It is the online interaction between people.

Discourse 2.0: The discourse available through Web 2.0 (Herring, 2011).

Epistemic stance: A stancetaker takes an epistemic stance through presenting oneself as "knowledgeable or ignorant" (Du Bois, 2007, p. 143).

Evaluation: It is "the process whereby a stancetaker orients to an object of stance and characterizes it as having some specific quality or value" (Du Bois, 2007, p. 143). Evaluation happens all the time as we take a stance. Sometimes it is direct, and others it is more implicit and needs to be inferred.

Evaluative act: It is the act of casting an evaluation (see the definition above) of the object of stance.

Hedges: Devices writers use like *possible*, *perhaps*, *might*, or other lexical or grammatical devices which introduce a lesser degree of commitment to one's assertion (Hyland, 2005). Hedges and boosters help writers determine where they would like their

words to fall on a scale ranging from fact to opinion. Hedges open a room for dialogue, dispute, and/or different interpretations (Hyland, 2005).

Positioning: It is the process of positioning oneself and others through taking stances along an affective and/or an epistemic scale(s) (Du Bois, 2007). It is "the act of situating a social actor with respect to responsibility for stance and for invoking sociocultural value" (p. 143).

Social media: It is a term that refers to internet websites which are built upon the interaction between participants and where the content available through them is published by people on that medium and is seen by the audience these people choose (Page et al., 2014).

Stance: It is an act whereby a stancetaker takes a position with regards to something (the object of stance), evaluates it, and thereby aligns or disaligns himself or herself with it (Du Bois, 2007). The stance taken by stancetakers are epistemic or affective.

Web 2.0: Web-based platforms which generate interaction by users (Herring, 2011).

Operational Definitions.

Accommodation: choosing to type in the same code as that used by the tweet author to whom the respondent is replying—as, for example, when a respondent replies in Arabic to an Arabic tweet. Non-accommodation occurs when a different code is used—as, for example, when a respondent replies in Arabic to an English tweet.

Arabized English (AE): This is the name given by the researcher to a newly-developed form of writing among polyglossic Egyptians where they write English in Arabic letters. The difference between Arabized English and transliteration is that Arabized English can be

written differently by different people, while transliteration has rules as to how it should be written.

Code: There are eight codes identified in this study, and the word code refers to any of them. These are Arabic, English, Arabizi, Arabized English, CS, "sounds, emoticons, and/or emojis," (SEE), no reply, and another language.

Code-switching (CS): In this study, code-switching refers to switching between any two of the following four codes: Arabic, English, Arabizi, and Arabized English even if that happened only on a word-level.

Computer-mediated Communication (CMC): It is any form of communication that is based on computer as a medium for it. In this study, Twitter is considered one.

Computer-mediated discourse (CMD): In this study, it refers to the online interaction between people on Twitter.

Egyptian bilinguals: In this study, this term refers to Egyptians who have weak, intermediate, or advanced command of Arabic and English languages.

Emoticons: They refer to typing a facial expression through letters and/or signs rather than word(s). Together with sounds and emojis, the three terms are considered one of the codes identified through the replies of the participants.

Emojis: Standardized expressive images (ideograms) used in electronic communication. Together with sounds and emoticons, the three terms are considered one of the codes identified through the replies of the participants.

Epistemic stance: In this study, a stance is identified as an epistemic one when the stancetaker presents himself or herself as knowledgeable or ignorant and when they present an opinion as if it were a fact. This occurs through stance markers, like hedges, boosters, and

verbs of knowledge, or through presenting information in fact-like statements through the simple present tense or through modals like *should*.

Evaluation: In this study, it occurs when the stancetaker presents an evaluation of the object of stance which can be something or someone. Markers for an evaluative act are adjectives.

Sounds: In this study, this term refers to instances of onomatopoeic reply—using a sound(s) in the tweet instead of word(s), like *oops* and *hmmm*. Together with emoticons and emojis, the three terms are considered one of the codes identified through the replies of the participants.

Stance: In this study, stance is the act whereby by the stancetakers take a position with regards to the content presented in the tweet they are responding to and/or its author, and thereby align or disalign themselves. The stance taken by the stancetakers was epistemic.

Stancetaker: In this study, the stancetaker is the respondent to the DCT. Each of the participants in the study is a stancetaker.

Polyglossia: Speaking more than one variety of a language along with another language.

CHAPTER TWO: LITERATURE REVIEW

Introduction

The present study examined the identity of polyglossic Egyptian users of Twitter. Understanding their identity occurred through an examination of the stance they took, their code-choice, and their linguistic accommodation. This chapter serves to provide the necessary background information to the present study through a review of the pertinent literature available. It is divided into nine sections: The first three sections focus on the language users, the relationship between their identities and the codes they use, the stance they take as they express themselves, and their choice of accommodation or lack of it and the reasons behind doing so. The fourth section discusses Discourse 2.0 which Twitter would naturally belong to. The fifth section presents a comparison between blogs and micro-blogs and explains why Twitter is considered a micro-blogging website, and the sixth explains the difference between diglossia and polyglossia. The seventh section sheds light on one specific code used by many Arabs on social media platforms *Arabizi*, and the eighth discusses the kinds of codes used on Twitter all over the world. The last one explains the mode of communication under study Twitter through explaining its conventions. This section is specifically directed to readers who do not know much about it.

Identity

Language is an indicator of one's identity and not just a means of communication between people (Bassiouney, 2009). Bassiouney (2010) states that for someone to create an identity, one has to express himself or herself. Studying people's use of code provides us with information about people's identities and their relationships with others (Holmes, 2013). Bassiouney (2010) explored the relationship between identity and code choice. She drew on Lakoff's (2006) definition of identity by stating that identity is not a simple shallow term; rather,

it is many overlapping things that define a human being. It is a person's history, background, beliefs, feelings, and many different things that are not the same for everyone. It is something that never stops developing as one grows. One cannot judge another's identity to be one thing or another (Bassiouney, 2009). Identity comes from a person's own beliefs and feelings of what and who that person is. Such beliefs are formed due to many factors—factors that are the interaction of years, history, and experiences—which eventually form what one believes oneself to be (Bassiouney, 2012).

Along the same lines, Paltridge (2012) maintained that people constitute different identities in different times depending not only on themselves but also on others, those with whom they interact. Two other aspects that define identity are the context and the purpose of discourse. Further, identity does not operate independently. It does, however, exist in relation to different aspects, like gender, ethnicity, race, sexuality, and other characteristics that define a human being. "Part of having a certain identity is that it is recognized by other people. Identity, thus, is a two-way construction" (Paltridge, 2012, p. 24). Moreover, texts have ideologies which underlie them. These ideologies carry biases, beliefs, ideas, and experiences, all of which can be political, historical, and/or social.

Bassiouney's (2010) article discusses how Egyptian women asserted their identities through switching between Standard Arabic (SA) and Egyptian Colloquial Arabic (ECA), yet she still made it very clear that the same applied to cases when women chose to switch between languages. Women, according to Bassiouney, aspired for power as a means of asserting their identity through feeling superior, which was why they chose to switch. Thus, a relationship exists between one's identity and one's choice of code.

Paltridge (2012) asserted this relationship between language and identity in these lines: "The way in which people display their identities includes the way they use language and the way they interact with people. Identities are no natural, however. They are constructed, in large part, through the use of discourse" (p. 24). In similar ways, people form online identities in order to feel connected with others, and in that sense, "micro-blogging can be seen as an on-going performance of identity" (Zappavigna, 2012). According to Paltridge, people also create online identities which could be different from their offline ones. These identities are created also through the use of language and also through the visual devices (Paltridge, 2012). This online identity exhibited through tweeting is what will be examined in this study with relation to all its constituting parts: code-choice, context, use of discourse, and interaction with others.

More interestingly, online blogging gives the opportunity for users to create new identities that are different from their real ones. Androutsopoulos (2006) thinks that this reflects the fragmentation of the post-modern identity. One can find the same idea explored in Elsadda (2010) where she mentions in her article that blogs provide anonymity for the blogger. Likewise, when one posts something to Twitter, or when a personal account is initially created, one can always stay anonymous. A person does not have to reveal their real character to people. Although many perceive cyberspace as a space for socialization, this is not always why people might want to create accounts on these forums of media. Wanting to hide is a conscious act one might do for many different reasons. Another way of thinking of the online identity is to think of it as one of the selves we try to project to others. Johnstone (2008) proposed that "current ways of understanding social identity and its relationship to discourse are rooted in the idea that the selves we present to others are changeable, strategic, and jointly constructed" (p. 155).

In one of her latest work about studying language and identity, Bassiouney (2014) says that to have access to a language is to have access to a resource. This resource is big enough that people have to constantly make choices which make them who they are. In other words, people construct an identity through the linguistic resources to which they have access. She summed up the relation between language and identity in the Egyptian society, explaining that:

> The clearest evidence of the immanent role of "access to resources" as a marker of identity is in the way that Egyptian public discourse utilizes language as a classification category … That is, in the projection of public discourse, the code that one chooses reflects directly on how one positions her or himself in relation to others: as an insider or an outsider, as an Egyptian or as a foreigner, as an Egyptian with no affiliation to Egypt, or as a loyal citizen, as a typical man in the street or as an Egyptian that does not share the same characteristics that unify Egyptians, and so on. In this scheme, for example, if one speaks Arabic, one is classified as Egyptian. However, since Egypt is a diglossic community, classification can also be dependent on which code is used and whether speakers switch between codes. (p. 41)

This is of relevance to the current study since, again, it sheds light on the code choice made by Egyptians and how it is directly related to their identity.

Stance

Bassiouney (2014) argues that stance-taking is a process in which people use language to give themselves an identity and try and impose another identity on others. In her explanation of stance, Jaffe (2009) postulates that the social meaning of linguistic forms resides not in categories like gender, age, and region, but rather in how identities are constructed through taking stances and creating alignments. In other words, Jaffe sets to investigate how stances are indexed either directly or indirectly through different linguistic forms.

Englebretson (2007) observed that stance is an evolving field of study among linguists. Englebretson added that stance could be quite a problematic term because its usage varied from researcher to another. In addition, some researchers preferred to use other terms to talk about the same phenomenon of stance. These terms included *evaluation* and *subjectivity*. Englebretson

summarized the extensive body of research done on the relationship between stance and the terms *evaluation*, *subjectivity*, and *interaction*. He noted that subjectivity was the broader term and idea of self-expression, while evaluation was focused subjectivity about someone or something.

Du Bois (2007) argued against the previous division of stance into types. He believed that stance was an act which involved a content and a context, an idea which was also stressed by Englebretson who argued that to best understand stance, the context must be looked into and understood. "Stance is a property of utterances, not of sentences, and utterances are inherently embedded in their dialogic contexts" (Du Bois, 2007, p. 148). Therefore, according to Du Bois, to be able to analyze a stance, we have to ask ourselves who the stancetaker is, what the object of stance is, and which stance the stancetaker is responding to—what he calls the *counterstance*. Asking these three questions is what helps us interpret stance. These three questions are the components of the dialogue from which stance can be understood.

Before Du Bois started his definition and explanation of stance, he explored the relationship between stance and dialogicality, intersubjectivity, and the social and linguistic frameworks invoked through stance. Because stance is taken based on a previous dialogue whether this is one taken on the same spot of the action going on or one that is based on previous old dialogues, a quick relationship is formed between stance and dialogicality. He defines intersubjectivity as the relation between one's subjectivity and another person's. Thus, intersubjectivity serves to "ground the sociocognitive aspects of stancetaking in dialogic interaction" (p. 141). The same idea is stressed by Englebretson (2007) who views the relationship between stance and intersubjectivity as a starting point for the relationship between stance and interaction. Lastly, taking a stance is a social as much as it is a linguistic act which

involves evaluation of something, and by taking a stance, other people take a stance based on the ones we took.

To explain stance, Du Bois (2007) explained three types of stance acts. The first type is evaluation. An evaluative act occurs when the stancetaker assigns a value or a quality to an object of stance. Evaluation happens all the time as we take stance, either by assertion or inference. The second type is subjectivity and positioning. Almost every act of positioning includes a degree of subjectivity by the stancetaker. This happens either when the stancetaker positions himself or herself on a scale of affective value (one's feelings towards something) or on a scale of epistemic value (one's knowledge about something). Du Bois then moved on to explaining the third type which is intersubjectivity and alignment. Intersubjectivity emerges when we observe the degree to which one's subjectivity (view or opinion) matches another's along an agreement/disagreement scale. Hence, aligning or disaligning with people occurs.

From the above explanation, Du Bois concludes that stance is not more than one act. Stance is a unified act analyzed by understanding its key components: evaluation, positioning, and alignment across a subjectivity/inter-subjectivity scale and thus forming the stance triangle. By taking a stance, people evaluate something and thereby position themselves with regards to it and either align or dislaign themselves with other(s).

Using Du Bois' (2007) framework, Haddington (2007) studied two methods of stancetaking in news interviews, namely positioning and alignment. Haddington identified that there were three practices of positioning done by interviewers (to position their interviewees) to which interviewees aligned themselves. The first positioning activity was asking hostile questions which the interviewees did not respond to directly; instead, they managed to align with the question and engage with it. The second activity was incorporating stances in questions

which prefer a subsequent stance in reply. Interviewees replied in hesitation marker to express their disagreement with the stance taken in the questions. After that, they showed more alignment through recycling part of the questions but with a negation in the middle. The last activity done by interviewers was using presupposition as a positioning technique to which interviewees replied with a stance marker followed by a linguistic recycling of a part of the question as done in the previous activity.

Academic writing is one of the registers where the analysis of stance has been carried out by many linguists. Hyland (2005) stated that academic writing has developed from a form of writing about a reality from which the writer is detached to a form in which the writer uses direct language that aims at convincing the readers of an argument. According to Silver (2003), academic writers intend to convince their readers by presenting valid, well-supported arguments. Therefore, academic writing is persuasive. Charles (2006) supports this view and states writers need to construct a stance through highlighting their individual contribution supported by evidence and research.

Stance was defined by Hyland (2005) as how writers expressed their ideas through language and whether they chose to stand out through their writing or whether they chose to hide through the language they were using. According to him, stance has three components: evidentiality, affect, and presence. Evidentiality is the writers' degree of commitment to their words and propositions; affect is the attitude, feelings, and perspective about them; and presence is the extent to which a writer chooses to stand out through the text (Hyland, 2005). These three components are not so remote from the stance acts identified by Du Bois (2007). Evidentiality is one of the methods whereby a person takes an epistemic stance, and affect is how one displays feelings about something and hence, take an affective stance. The difference between Hyland's

and Du Bois' propositions lie in the fact that Hyland's is about academic writing which is a different register with its own defining characteristics.

His analysis of stance led Hyland to postulate that writers have recourse to the following four components: hedges, boosters, attitude markers, and self-mention. Following is an explanation of each. Hedges are devices writers use like *possible*, *perhaps*, *might*, or other lexical or grammatical devices which introduce a lesser degree of commitment to one's assertion (Hyland, 2005). Boosters are devices like *certainly* and *obviously* which reflect how positive a writer is about the information presented (Hyland, 2005). Hedges and boosters help writers determine where they would like their words to fall on a scale ranging from fact to opinion. Hedges and boosters are given other names by Silver (2003) according to their function. Hedges are called *downtoners*, *downgraders*, or *understaters*, while boosters are called *overstaters*, *intensifiers*, or *emphasizers*. Reasons why writers might choose one of these devices (over another) could vary. While hedges help open a room for dialogue, dispute, and/or different interpretations, boosters shut such a space and invite readers instead to a more reliance on the writer's words. Both hedges and boosters were examined in this study to understand the degree of commitment of stancetakers to their own words and thus reach an understanding about their stance and their identities.

Attitude markers convey a writer's attitude to a text (Hyland, 2005). This attitude could be that of surprise, fear, (dis)agreement, or any other attitude a writer wishes to convey through the text. Hyland (2005) describes the devices used to express attitude as follow: "While attitude is expressed throughout a text by the use of subordination, comparatives, progressive particles, punctuation, text location, and so on, it is most explicitly signaled by attitude verbs (e.g. agree,

prefer), sentence adverbs (unfortunately, hopefully), and adjectives (appropriate, logical, remarkable)" (p. 180).

The last device for expressing stance according to Hyland is self-mention which denotes the presence of writers in their words. Different writers in different fields resort to doing so in different ways and for different reasons. In some cases, it is better to be absent and to present ideas in a rather general manner, and in other cases, it is better for writers to stand up and to clarify that they are the voice behind such a perspective or opinion.

Hunston (2007) discussed the importance of using corpus in analyzing stance. Corpus Linguistics is a field of study that has allowed linguists and researchers to identify both quantitatively and qualitatively the frequency and function of use words and/or phrases (Hunston, 2007). It offers context and numbers which assist researchers in reaching their findings about the linguistic item under investigation. Hunston concluded that in analyzing stance, corpora can be used to quantify stance markers and to offer context to the words which eventually give qualitative work a more reliable meaningful worth. Context is what leads to the potential of always referring back to how structures were used and why, in order to eventually be able to get their functions. However, qualitative analysis must be used alongside the quantitative analysis to make sure the results reached are accounted for and are well-explained. Although this study is not done on corpora, the data collected will be analyzed mostly qualitatively but will also include quantification of some data to increase the study's reliability and meaningfulness.

In a study about how speakers and writers take stances through the use of adverbials, Conrad and Biber (2000) investigate epistemic stance adverbials, which indicate how certain the speaker or writer is, attitudinal stance adverbials, which determine a person's attitude or feelings about what is being said or written, and style adverbials, which determine the comments on what

is said or written stances. This was done through an analysis of a corpus of news reportage, a conversation, and an academic writing text. The results showed that stance adverbials were used in conversation almost twice as frequently as they were used in the written registers. Further, epistemic stance adverbials were the most frequently used in all genres of the corpus. In the academic register and the news reportage, the second most used stance adverbial is the attitude followed by the style adverbials. However, in the spoken register, style adverbials were used more frequently than attitude adverbials.

Johnstone (2007) explored in her study how stance-taking linked between linguistic forms and social identities. This was done through an analysis of a conversation between Pittsburghers where she identified how their dialect identity was linked to their local identity. In other words, being identified as a Pittsburgher was linked to speaking Pittsburghese. Johnstone reached this finding through examining direct and indirect methods of taking stance through claiming authority. Her study reflects the link between code, stance, and identity which is what this study aimed to investigate about polyglossic Egyptians.

In another study, about spoken Finnish, Rauniomaa (2007) analyzed the use of two stance markers used as hedges. She drew her data from a corpus of spoken Finnish, and found that both two stance markers were often linked to assessment. The functions they performed were to show disagreement in a second assessment, to mark transition to a first assessment by another speaker, and to mark transition to a first assessment by the same speaker.

Keisanen (2007) studied how stance taking was done in interaction to show how interlocutors (dis)align with each other. Data were taken from the Santa Barbara Corpus of Spoken American English (SBCSAE) where Keisanen examined negative yes/no interrogatives and tag questions. Results showed that interrogatives were used as disaligning actions to

challenge what was previously said or the stance taken by a previous speaker. These disaligning actions embodied epistemic as well as affective stances, and they resulted in delay in response and repair initiators, which meant that the respondents realized the interrogatives were used as markers of disalignment.

In a yet another study exploring the relation between the use of codes and taking a stance, Bassiouney (2015) collected the data of 17 celebrities in interviews, some of whom performed (in their performances as actors or singers) in ECA and some did not. Bassiouney concluded that Egyptian media-makers were trying to force ECA into being the norm and the standard code, and the fact that it no longer is was the very reason why they were trying to do this. However, celebrities those who performed in ECA and those who did not took different stances toward this. Some of them spoke totally in ECA, some resorted to code-switching in their interviews, some refused to speak in ECA marking a stance of pride in their own national identity. Further some celebrities went so far as to claim that they had become Egyptians by choice while others stopped at admiring and expressing love for Egypt. She also concluded that North African artists were the fastest to acquire ECA and even to conduct full interviews in it. Syrians followed by Lebanese celebrities code-switched between Levantine and ECA the least. In other words, they used less ECA and some of them even did not use ECA at all.

Accommodation

Giles and Powesland (1997) clarified that the theory of accommodation was based on social psychological research which stated that by reducing differences, people became more favorable to each other. When people accommodate their speech, it results in being more similar to the person they are speaking to. The same idea was explained by Johnstone (2008) who argued that an individual's style of speaking is shaped by the styles of those to whom he or she is

speaking. An individual may either speak in similar ways, signaling similarity, or in different ways, signaling difference and unbelonging. Shifts in style occur according to the audience and also according to the topic, either consciously or unconsciously. Accommodation, according to Johnstone, occurs through code choice, accents, and also though "stylistic and rhetorical decisions on all levels" (p. 148).

Giles and Powesland (1997) offer two possible reasons why people might resort to accommodation: their desire to gain social approval or to be better understood by making sure their message is clear to the receiver. Giles and Powesland also, however, explain that in real life interactions, there are several underlying motives for (non)accommodation and effects which speakers might aim at through (non)accommodation.

Giles & Powesland extended Heider's (1958) attribution theory—of attributing motives to an act—to accommodation. They explained three factors that need to be thought of when thinking of the motive behind accommodation or lack thereof. These are the speakers' ability to accommodate, their effort to reduce the dissimilarities between them and the hearer, and the pressure which might be exerted on them to accommodate. Voluntarily making the effort to accommodate is set in contrast to being pressured to do so, which affects the listener's perception of the speaker and of the act of accommodation.

Discourse 2.0

Herring (2007) defines Computer-mediated Discourse (CMD) as the interaction between people over the internet through computers or mobile phones. At present, there are many gadgets which allow for internet access and it would be better to replace the term *CMD* with the term *Discourse 2.0.* Herring (2011) defined Discourse 2.0 as the discourse available through Web 2.0

which she defined as web-based platforms which generate interaction by users. In her opinion, Discourse 2.0 “offers a rich field of investigation for discourse analysts" (p. 21).

Tannen (2013) argued that when people communicate, they do not just communicate the literal messages, but they also communicate metamessages, such as how the message is intended by the speaker, how it is interpreted by the hearer, and why this message was said in this context, this way, at that time. The metamessages observed on new media by Tannen were markers of (un)enthusiasm, markers of (in)directness, turn-taking, and medium choice. Tannen argued that these aspects were in some ways similar to and in others not very similar to aspects of conversation. Some of these metamessages will be explored in this study as part of the Discourse 2.0 on Twitter.

Herring (2011) proposed a three-part classification scheme of Discourse 2.0. These were the familiar, the reconfigured, and the emergent aspects of the discourse. The familiar aspects of the discourse include the use of text as a channel of communication, "nonstandard typography and orthography, code switching, gender differences, flaming, and email hoaxes and scams" (p. 8). The reconfigured or the reshaped aspects of the discourse which include "personal status messages, quoting others' messages, small stories, and customized advertising spam … as well as configurations of such familiar phenomena as topical coherence, turn-taking, threading, and intertextuality" (p. 10). Finally, the emergent aspects are those which are entirely new and unprecedented, and they include "the dynamic collaborative discourse that takes place on wikis, along with conversational video exchanges, conversational exchanges via image texts, and multimodal conversation more generally" (p. 14). This emergent aspect was created and used by internet users who might not even know one another. Twitter is a medium which includes this

three-part classification scheme. However, the familiar and the reconfigured will be the only ones explored in this study through examining the code use and replying to a tweet author.

Herring (2007) presented a method of classification for Discourse 2.0 to help linguists use and conduct more research on it. In doing so, she followed the example of how Discourse Analysis (written and spoken discourse) was developed. In order to make it easier for understanding, analysis, and comparison, Discourse Analysis (DA) is based on classifying the type of discourse into several criteria including "modality, number of discourse participants, text type or discourse, and genre or register" (p. 4-5). Herring attempted to do the same with Computer-mediated Discourse Analysis (CMDA).

Herring (2007) explains that when it comes to CMD, two main influences are at play: the social (or situation) influence and the medium (or technological) influence, and hence the term *social media*. Under each of these influences, a number of facets (or categories) are presented. The social set describes the features surrounding the human participant, like the number of participants, the purpose of communication, the topic, and the language used. The medium set describes the technological features of the mode of communication, like the type of message, its size, and how it is transmitted.

Table 2.1 explains Herring (2007) medium and situation factors and their application to tweeting. It is quite necessary to mention here that typing a private message through chatting on Twitter (through what is called Direct Message (DM)) is different from typing a message through posting a tweet. Since this study is concerned with tweets and not DMs, the characteristics of the medium and social factors in Table 2.1 will only apply to typing tweets.

Table 2.1

Herring's (2007) Medium and Situation Factors and their Application to Tweets

		Factors	Details	Application on Tweets
Medium Set	1	Synchronicity	Synchronous vs. Asynchronous	Asynchronous
	2	Message transmission	One-way vs. Two-way	One-way
	3	Persistence of transcript	Ephemeral vs. Archived	Archived
	4	Size of message	Amount of text per message	140 characters
	5	Channels of communication	Text, visual, audio	Words, image, sounds, video, link
	6	Anonymous messaging	Presentation of participants' identity	High
	7	Privacy messaging	Having private conversations with other users	N/A
	8	Filtering	Ignoring messages	N/A
	9	Quoting	Quoting a message in the reply	Replying, retweeting, or quoting a retweet
	10	Message format	Order of messages	Sequenced with time
Social Set	11	Participant structure	• Number of participants • Privacy settings • Anonymity	• Endless number of participants • Public, semi-private, private • High
	12	Participant characteristics	Self-identification (background, skills, ideologies, attitudes, etc.)	Topic, attitude, following
	13	Purpose	Goals of interaction	Sharing, forming relationships/hostilities, exploring, etc.
	14	Topic	Content/subject	News, jokes, gossip, ideas, feelings, etc.
	15	Tone	Degree of formality and seriousness	Flexible
	16	Norms	Norms of organization and norms of language	No commitment
	17	Code	Language variety	Flexible

Herring (2007) acknowledged that her list was not inclusive and that new factors could be added or subdivided and that existing ones could go out of use, especially with the incredible pace with which new technologies develop. An explanation of each of these factors will follow with the same numbers it was given in the table:

1. Halim and Maros (2014) made a distinction between two modes of communication: synchronous and asynchronous modes. Synchronous communication is the type of communication when the reply is expected to take place with no delay, probably without investing much time on thinking, like what happens on an internet chat where the participants are online and are having a conversation. Asynchronous communication is the type of communication when people are not necessarily expected to be online and reply instantly but are expected to take their time to reply. To send an e-mail, for example, the sender and the recipient do not have to be logged in at the same time. Instead, the sender can send the e-mail which is then stored and the receiver finds it when is logged in (Herring, 2007). The same applies to typing a post (or a tweet in the case of Twitter). Asynchronous modes of communication can also form threads of discussion on social media websites which serve to make these discussions easier (Hansen, Shneiderman, & Smith, 2011). Some websites, like Facebook and Twitter, fall under both types of communication. However, for the sake of this study, the asynchronous type shall be the only one examined through tweets.
2. According to Herring (2007), two-way transmission of messages is when the recipient views the message that the sender is typing while it is being typed or when an indication that a message is being typed is shown on the recipient's screen. One-way transmission is when the recipient has no clue that a message is being typed except when it is received.
3. A message is ephemeral if it remains for a short time on the system. It is archived if it is stored there and a person can get back to it whenever is needed (Herring, 2007). Tweets are archived and any Twitter user can find all their tweets. The difficulty lies when a

person has hundreds or thousands (or more) of tweets, then they have to scroll down for a very long time to reread their first tweets, but they can eventually be found.

4. Some forms of communication apply limits to the number of characters per message, and hence control its size, while others do not (Herring, 2007). One of Twitter's defining characteristics is that a tweet's size is very small, limited to 140 characters.
5. A channel of communication is what form the message takes. A message can be text, video, audio, image, or conference (Herring, 2007). Through a tweet, one can type a text and/or share a video, an audio clip, an image, or a link to another website.
6. Some websites offer their users easy access with no identity verification. Others require some sort of verification, like e-mail verification. This contributes to the degree by which a user stays anonymous (or not). People can be totally anonymous on Twitter. As a matter of fact, Twitter is not the ideal means to truly know a person, because they can easily adopt a totally different identity than the ones they project in reality. To sign up for Twitter, users provide their full name, e-mail address, password, and username. These details can easily be formed for a fake account whilst the identity of the real user is kept anonymous. Some accounts though are "verified," particularly for public figures and celebrities.
7. Private messaging refers to the ability to carry on private conversations (Herring, 2007). The only way to have a private conversation with another Twitter user is through the DM function. Tagging a user in a tweet is a way of sending a message to a specific user, but it will be seen by other followers so the characteristic of privacy will be violated.

8. Filtering is the ability to "ignore messages from another user" (Herring, 2007, p. 16). Ignoring tweets of a Twitter user can be done through blocking or unfollowing that person. However, if one is tagged in a tweet even without following its sender, one will still get that tweet.
9. Some websites facilitate quoting a previous message in the reply and others do not (Herring, 2007). Twitter allows quoting tweets in several ways. First, upon selecting the first tweet, all the replies to it are sequenced underneath it. Second, one can retweet someone's tweet. This means that the whole message is copied while making reference to the original sender of the tweet. Third, a person can quote a retweet which means that the whole message is copied while making reference to the original sender along with facilitating a space for typing a comment by the person who is retweeting the tweet. These are all possibilities facilitated by the system itself (how Twitter works).
10. Message format is the order by which messages appear and how they are presented (Herring, 2007). As mentioned in the point before, two or more can reply to a tweet and the order of their replies will be displayed with the newer replies appearing by scrolling down the page.
11. Participant structure refers to three things: first, there is the number of participants which can be endless on Twitter. Second, there is a privacy setting which concerns the whole account ranging on a scale from private to public. The space of one's profile is private, semi-private (or semi-public), or public according to each user's preference. Tweets can either be public or restricted to approved followers only. Third, there is a degree of anonymity by which a person chooses to appear. This depends on the users and how much information they choose to show about themselves. It is more or less the

same as point number 6 (in the medium factors) but more from the participants' perspective and how they want to be viewed by other users.

12. Participants' characteristics are their ideologies, gender, attitude, beliefs, skills, experience, or anything that could have an effect on the kind of interaction that might take place between participants (Herring, 2007). It is the information which users choose to display in front of others through their profile page. On Twitter, this information is a name and two pictures they choose for themselves and a 140-character limit to write anything that could identify them. This is why there is a wide room for staying anonymous on Twitter quite easily. The importance of such information is that it contributes to the identity a person tries to project. This could have an effect on determining the kind of people a user might connect with, on the topic of users' tweets, on their attitude towards each other, and on choosing who to follow and who not to.
13. Purpose is the goal(s) that participants have behind their interaction (Herring, 2007). The purposes of interaction on Twitter are as many as the purposes of interaction people might have in face-to-face interactions. There is not a definite number; they can vary widely. People's aims behind interaction on Twitter can be, for example, to form relationships, to keep up with news, to explore ideas, to advertise products, to express political opinions, or to comment on news or current events..
14. Page, Barton, Unger, and Zappavigna (2014) suggest that social media is the means through which people exchange information, feelings, news, ideas, jokes. These form the content of interaction between people in real life as well as on Twitter.
15. Tone is what users adopt online on any medium. It can range from formal to informal, from serious to light (Herring, 2007). Determining the tone depends on all of the above

factors, and different people can adopt different tones. This is the case with Twitter users.

16. Norms refer to conventions of communicating online. These apply to language, as in the use of abbreviations, acronyms, and jokes, for examples. They also apply to practices within a group, like actions taken against violators by group admins or leaders. On Twitter, these conventions are somewhat flexible because Twitter has a more spontaneous nature and has no groups that could be expected to maintain certain norms in the sense that Facebook, for example, does.
17. Code refers to the language or the variety used by the participants (Herring, 2007). The codes people use to tweet can be any code and can be a mix of codes too. This depends on a user's personal choice which can be affected by a number of factors including the participants involved, the topic under discussion, and the identity they wish to project. This point is of particular importance to the current study since it is the point around which the research is based.

Blogs vs. Micro-blogs

There are many forms of online writing. A quick comparison between two of them can be very beneficial here, namely blogs and micro-blogs. Whereas a blog has longer posts, a micro-blog has much shorter ones. To be a blogger, one does not have to be a famous renowned writer, but has at least to know how to own and manipulate the word, to hold on to an argument for long enough and to express and elaborate on a bigger scale. To micro-blog, however, one can express things in very little words. A person is not expected to delve into deeper explanations or arguments over anything. Java, Finin, Song, and Tseng (2007) observe two basic differences between blogs and micro-blogs. The first is that micro-blogs are considered faster modes of

communication where the time invested thinking about what to write is much less than that spent thinking about the content of a blog post. The second concerns the frequency of posting. Whereas bloggers are expected to update their blog "once every few days," a micro-blogger may often update more than once daily (p. 2).

Liu et al. (2012) describe Twitter as an online micro-blogging service, and many researchers think of Twitter as synonomous to a micro-blog. Indeed, it is a prototypical micro-blog. One of the defining aspects of Twitter is that its users understand that they have a character limit per post. This fact makes the data under investigation interesting, because language use therefore includes being careful about the wording used to communicate the message that users want to express (Zappavigna, 2012). They have a very limited number of characters in order to do so. Micro-blogs “create a live stream of bite-sized information nuggets” (Hansen et al., 2011, p. 23).

Elsadda (2010) investigated how online blogs affected Arabic Literature through a study of three blogs by three different Arab women which gained wide fame to the extent that they were published. Talking about blogs, Elsadda (2010) explained:

> Literary blogs are many things at the same time. They are forums for consciousness raising, social transformation and political mobilization. They are diaries, narratives of the self that are no longer locked up in drawers but made available to an audience. They probe the intimate secrets of the self, which is on display, even if under a false name. They are also messages or letters sent out to an imagined virtual audience. In actual fact, literary blogs defy generic classification: they are invariably a mélange of diaries, memoirs, autobiographical stories, to-do shopping lists, political manifestos, reflections, epistolary narratives, short stories and novels. (p. 328)

Although this is said specifically of blogs, many of these do apply to micro-blogs. People tweet about different things for different reasons.

Thinking about the reasons why people choose to and continue to tweet were the questions which Java et al. (2007) pondered. Through their analysis of a corpus of 1,348,543 tweets of 76,177 users, Java et al. classified the uses of Twitter as follow: a) posting about the daily routines, what people are doing during their day; b) having conversations with people they know on Twitter, through name-tagging; c) sharing information, especially through adding links of websites; and d) commenting and expressing opinions on current news. They were also able through their study to categorize Twitter users into three categories: a) the user who is a source of information. This user has many followers due to the valuable information that person posts; b) the user who is a seeker of information. This user rarely tweets but uses Twitter mainly as a source of information through following what others to read their posts. Finally, c) users who are friends. Those users have each other on the following list due to being friends, family, or co-workers. This category, at least now, does not fit much into the Twitter environment, since most users follow and are followed by some people they know and also many other users whom they do not personally know.

Krishnamurthy, Gill, and Arlitt (2008) offered a somewhat different but closely related categorization of Twitter users. With a corpus of close to 100,000 users, Krishnamurthy et al. categorized users into three types. The first consists of those who have more followers than people they follow. Many of them tweet a lot. These are celebrities, companies, politicians, and people whom they call *tweet broadcasters*. The second type consists of those who have a closely similar number of followers and followees. The number of people on both following lists increases with the number of tweets. People in this group "tend to exhibit reciprocity in their

relationships" (p. 2). The last type consists of those users who have a much bigger number of followees and are not followed by nearly as many people. These are the users who follow as many people as they can in hope that they would be followed back, and they do not tweet a lot. Although these results were reached in 2008, personal observations do support these results. Yet, the last type of users could also be those who use Twitter as a source of information. They may not necessarily be interested in tweeting but follow accounts to read tweets posted on them.

Diglossia vs. Polyglossia

This section aims at explaining the codes used by Egyptian bilinguals whose tweets will be analyzed in the current study. This will be done by explaining two important terms: diglossia and polyglossia. Ferguson (1959) defined Diglossia as the presence of a high (H), often regarded as superior, and a low (L), often regarded as inferior, varieties of the same language. The former is used in formal occasions, as in news reporting, and the latter is used in informal occasions, as in everyday conversations with friends and family. The H variety was called by Ferguson the *superposed variety* since it was not native to its speakers but rather learned through formal education, and the L variety was called the *regional dialect* since it was the native one which speakers did not need to learn through a formal kind of education. Another difference between both varieties resides in the fact that H has dictionaries, vocabulary, grammar, and pronunciation written and described, whereas L does not or has a very limited body of formal or published description. Ferguson stresses that diglossia is not a stage at a point in time during a process that a society's language is going through, like the process of standardizing a language; rather, it represents the actual use of language in that society.

Over the years, following its definition by Ferguson, diglossia was extended to include the presence of two or more codes used whether these are of related varieties (of the same

language) or of different languages. Later, diglossia gained a refined definition by Fishman (1967) who explained that these codes can be "functionally differentiated language varieties of whatever kind" (p. 30).

In an attempt to explain diglossia and its relation to bilingualism, Fishman (1967) draws a comparison between four speech communities: a) one where diglossia exists along with bilingualism; b) one where diglossia exists without bilingualism; c) one where bilingualism exists without diglossia; and d) one where neither of them exists. Before drawing on each of these communities, it is worthwhile to explain how Fishman defined diglossia and bilingualism. According to him, bilingualism is an attribute of a linguistic behavior of individuals, and diglossia is an attribute of the organization of codes on a societal and communal level.

A description of the four communities offered by Fishman follows. The first community is characterized by there being separate roles for each of the languages used in the same society triggered by a difference in the community's social levels. These roles are the result of a relation of power in that community and are distinctly understood in terms of when, why, and with whom they should be used. The second community is one where two codes (whether varieties of the same language or two different languages) are used in the same society, each has separate different role(s) but one is restricted to a class of that society and the other is restricted to the other class. In such case, each of the classes speaks only in its code and negotiation between both classes occurs through interpreters or translators. In the third community, codes are not restricted to a certain class or group of people rather than another. Instead, the whole community speaks and is able to communicate in all the codes. Members of the fourth community interact together with no differentiated registers or codes. Such communities are very hard to find, if they do exist at all.

Platt (1977) calls the first speech community which Fishman described *polyglossic*. Polyglossia is defined by Platt as diglossia coupled with bi-(or multi-)lingualism. He draws his evidence from the presence of such communities in Singapore and Malaysia, specifically "the English-educated Chinese speech communities in these two countries" (p. 363). Platt clarified that their speech repertoire consisted of the following codes[1]:

(1) 'Native' southern Chinese dialect.

(2) Other southern Chinese dialect(s)- particularly if own is not the dominant dialect of the region.

(3) Little or no Mandarin.

(4) Some formal Malaysian English.

(5) Some colloquial Malaysian English.

(6) Some knowledge of Bahasa Malaysia (the official standard Malay of Malaysia).

(7) Bazaar Malay (Bahasa Pasar).[2] (p. 365)

Platt calls this an example of a complex polyglossia where several H, Medium (M), and L codes are used. Whereas Platt added an M here as a third variety, other linguists, like Hurreiz (1975) applying his ideas to Sudan, think of varieties as a continuum rather than as specific stages.

Applying the above to the context of this study, Egypt would be considered a polyglossic community, one where more than one variety of a language (Arabic) exists along with another language (English). Although not all Egyptians are polyglossic, this study will examine a section of polyglossic language speakers in Egypt.

---

[1] Platt used the term *sub-codes* to refer to dialects or varieties of the same language and used the term *codes* to refer to "genetically distinct languages" (Platt, 1977). However, in this study the term *code* is used to refer to any language, dialect, or variety spoken in any country.

[2] For a more detailed background and explanation, read Platt's (1977) "A Model for Polyglossia and Multilingualism (With Special Reference to Singapore and Malaysia)".

Arabizi

Abu Elhij'a (2012) contemplated that the advent of the internet and the social media websites have helped young speakers of Arabic to develop an electronic form of writing in which people write the language they speak in order to use it in their electronic media interactions. The development of electronic writing has had a strong effect on diglossic languages like Arabic. The kind of electronic writing that was invented by Arabic speakers was called, by Abu Elhij'a, *Electronic Amiyya* (EA). She defined it as a system of writing spoken Arabic using Latin script. EA was the means by which people, especially young adults, were able to communicate freely using their own codes without having to switch to a more perceived formal form of writing which they do not use in their everyday communications (Abu Elhij'a, 2012). The same system was explained by Yaghan (2008) to be an Arabic language system which is written using Latin characters. Yaghan, however, gave it a different name *Arabizi*. As introduced by Yaghan (2008), Arabizi is a term which results from the combination of the two Arabic words "'arabi' (Arabic) and 'engliszi' (English)" (p. 39).

Yaghan (2008) asked some students (no specifications about them was mentioned) about their reasons for using Arabizi. Many different answers were given in reply to this question including: a) Arabic typography was not supported by early computing and mobile phone devices and technology; b) Arabic was more related in the minds of some as a language to write Classical Arabic (CA) in and not the everyday language; c) Arabizi supports the upper and the lower cases of English which conveys emotions more than Arabic letters do; d) Latin characters offer the phonological variation which Arabic characters do not which in turn serves to clarify the pronunciation based on individual or dialect differences; e) the number of characters allowed in typing a message in English letters is more than those allowed in Arabic letters (however, at

present, this is no longer the case); f) Arabizi allows for an easier way of code-switching between Arabic and English; and g) Arabizi is a flexible code that is not formalized or taught and, hence, is free of errors.

Along similar lines, Muhammed (2013) provided three reasons for the creation of Arabizi in internet communication. The first one was that the beginner users were North Americans. The second was that scientists used the Latin alphabet as the code for developing computers. This code is the American Standard Code for Information Interchange (ASCII). Hence, Arabizi was the means to overcome limitations caused by ASCII and the state of technological development at the time. The third reason was that English served as the universal language that connected different users of the internet all over the world.

Concerning the reasons why many people use Arabizi, a hashtag "#ليه_بتتكلم_فرانكو"—which can literally be translated into #WhyDoYouSpeakFranco[3]—started trending (using the Twitter language) on February 18, 2016. A quick run through the tweets people typed for this hashtag showed that they had different reasons for doing this. Some of them thought that it made them seem cool, and others expressed that it was easier in typing. Some people tweeted that they did not use it, and that Arabic was the code they preferred to use. Others expressed their total discontent about this code saying that people should either type in English or in Arabic. Still others voiced their concern about the future of the Arabic language and expressed that this tendency to write Arabic in English letters is actually posing a threat to their culture, country, history, and even religion. One worth noting thing was that tweets with this hastag were mainly written in Arabic, few in Arabizi, and very few code-switched between Arabic and English.

[3] *Franco* is the ECA word used to describe writing Arabic in English letters. In other words, it is the term Egyptians use colloquially to describe Arabizi.

In the English-speaking context, Merchant (2001) explored the language of teenage girls in online modes of communication and the changes spotted in their language use over these CMC modes. He explained that the long-held belief about the difference between written and spoken languages is that spoken language is thought of as temporary and taking place only in face-to-face communication, whereas written language is thought of as communication that is meant to last over time. CMC modes have enabled sound, text, and images to be added to online conversations, and thus moved the features of face-to-face interactions to online-based interactions (Merchant, 2001). Merchant attributed the change in language to teenagers who have become innovators of a new language written online which is characterized by lack of spelling and grammatical accuracy and the use of jargon and abbreviations. These characteristics facilitated saving time through quick typing and also added a sense of informality to make written interactions closer to informal day-to-day conversations. The same characteristics apply to Arabizi.

Arabizi conveys the informal language used in speaking through writing. This is facilitated by the characteristics of Arabizi which are: the use of Latin characters, the development of a system of abbreviations by Internet users, phonetic spelling, the use of emoticons through non-alphabetic characters, like the colon and the parentheses, forming acronyms through combining the initials from different words, and combining letters and numbers to form some words (Merchant, 2001).

One of the very few related studies which examined code-switching on Twitter in Egypt was that of Kosoff (2014). The codes Kosoff (2014) codified were English, Modern Standard Arabic (MSA), Egyptian Colloquial Arabic (ECA), Arabizi (MSA), Arabizi (ECA), and Code-switching (CS) between any two of these codes. Kosoff examined the tweets of ten Twitter users

over a period of time and counted the frequency of use of each of these codes in their tweets. The quantitative results presented in the study showed that more than 90% of the tweets of 5 (out of 10) participants of her study were in English, 4 users (of these 5) and another user code-switched between English and Arabizi (ECA), 3 users used MSA and ECA more frequently than any other code, and one user made a combination of all of these different codes.

Kosoff found that the code-choice of each of the users depended on the kind of audience they were targeting. For example, those who used English and Arabizi (ECA) targeted the upper classes who were likely to have been well-educated enough to understand the message and the Roman characters in the tweets and who were probably wealthier customers. Using code-switching between MSA and English was intended to reach as much audience as possible. Using ECA was seen as an attempt to reach a wide Egyptian audience. MSA was viewed as a means to maintain a certain position taken by its user in society. Arabizi MSA was rarely used due to the association of Arabizi with informal writing and the association of MSA with formal levels of writing. Hence, users' tweets were taken "to describe the socioeconomic and educational background of the Twitter user's target audience" (Kosoff, 2014, p. 98).

Language Use on Twitter

Many studies examined the most commonly used languages on Twitter. Krishnamurthy et al. (2008) concluded from their corpus that the top 11 countries in which Twitter was used were the USA, Japan, Germany, UK, Brazil, Holland, France, Spain, Belgium, Canada, and Italy, with these countries comprising around 50% of the countries in their dataset of tweets. Burger, Henderson, Kim, and Zarrella (2011) examined the top 13 languages used on Twitter. These were English, Portuguese, Spanish, Indonesian, Malay, German, Chinese, Japanese, French, Dutch, Swedish, Filipino, and Italian.

I shall draw upon Hong et al.'s (2011) study which aimed at examining the use of different languages on Twitter and how speakers of different languages behaved on it. Hong et al. (2011) examined a corpus of around 62 million tweets through which they identified the top 10 languages used and the use of URLs, hashtags, retweets, mentions, and replies by different language tweet authors. One of the gaps indicated through their study was that most of the studies done on Twitter only examined English tweets assuming that English was the universal language and that the conventions used in English tweets applied to all other languages on Twitter. Therefore, and through their study, they set to prove this wrong.

One hundred and four languages were used in the data of 62,556,331 tweets used in the study (Hong et al., 2011). The top 10 languages used were English, Japanese, Portuguese, Indonesian, Spanish, Dutch, Korean, French, German, and Malay, comprising 95.6% of tweets in their corpus, meaning that 4.4% of the tweets examined were written in the remaining 94 languages. 51.1% of the tweets were in English. The results also showed that Twitter users of all languages retweet, use URLs, hashtags, mentions, and replies, but the percentage of usage of each of them differed according to the language. This means that cross-cultural differences do appear even in the way different language users use Twitter as a social media website.

In their analysis of the reasons why the use of different Twitter conventions differ according to the language community, Hong et al. (2011) suggest some possible explanations such as cultural differences, how long different users or different communities had been using Twitter, and whether some of these users were bilinguals who might have affected the way these conventions were used in a given community. Their study is relevant to the current study which examines users who are bilinguals with an access to English, the top used code on Twitter, and Arabic, one of the codes which, statistically-speaking, is rarely used on Twitter.

Twitter

Launched in 2006, Twitter is primarily a social network site which also allows people to create their own individual profiles which they can keep hidden from other users. What users post on Twitter is called a *tweet*. Users tweet about anything about which they want to express a feeling or opinion. Twitter users also follow each other's tweets, thus forming a web of followers and followees. Like most online blogging sites, Twitter has some conventions of its own. Among these conventions are usernames, name tags, replies, hashtags, and retweets. Following West and Trester (2013), Twitter conventions and terminology are summarized in Table 2.2 followed by a detailed explanation of each.

Table 2.2

A Summary of Twitter Conventions and Terminology

Twitter Term	Description
Username	Formed by the e-mail handles of users
Name tag	The sign @ followed by a person's username and is used to tag users in tweets
Reply	Direct response to someone's tweet and appears underneath the tweet
Hashtag	The pound sign # followed by word(s) and is used to define the topic or comment on it
Retweet	Including an old tweet in a new one with or without adding a comment

Usernames on Twitter are the handles of the e-mails of users (i.e. the characters before the @ of the e-mail address). Name tags are formed with the sign @ followed by a person's username. Hong, Convertino and Chi (2011) note that nametags are used in tweets to include the tagged people in the conversation or to grab their attention, which is one of the ways in which Twitter facilitates communication between people, essentially making Twitter a medium for social interaction. Similarly, a reply is "a specific form of mention with @username appearing at the beginning of the tweet, is a tweet responding to a previous message" (Hong et al., 2011, p. 520). On Twitter, a user replies to another user whom they know or get to know over Twitter. It

could also be a celebrity or a public figure or just someone who posted a tweet they wanted to reply to. Hence, replying is a form of maintaining strong social ties suggesting using Twitter as a medium of social communication rather than a medium of getting information or merely expressing oneself (Hong et al., 2011).

Doyle (2012) defines hashtags as the words written with the pound sign (#) preceding them, like *#hungry*. Hashtags consisting of more than one word are either written with no intervening spaces or with intervening underscores (_). Doyle explains two ways of using a hashtag. The first is the hashtag as a labeling category. This was the first use which hashtags were created for. In this sense, a hashtag is added to classify the tweet as belonging under one topic or another. An example from a random tweet taken from Twitter reads: "'Majority of these men don't want help. They believe they own these women, and they will beat these women.' bit.ly/1C4yvcN #abuse" The link in the tweet is to a newspaper article about abused women. The hashtag *#abuse* in this tweet is used to classify the tweet as falling under the topic *abuse*. Through this explanation of this use of hashtags, Hong et al. (2011) posit that hashtags can help categorize the tweets about different topics as they are worded in the hashtag. The second use is the hashtag as a paralinguistic marker which Doyle (2012) calls a *meta-hashtag*. Use of this hashtag is slightly different in the sense that its main aim is to add a comment to the body presented in the tweet (Doyle, 2012). An example from another random tweet reads: "Drove all the way to Woodland Hills to visit my parents at work and wasn't even allowed to go in the building #sweet." The hashtag *#sweet* in this tweet is a sarcastic comment added by the writer to comment on the situation at hand.

Retweeting was defined by Herring (2013) as the “inclusion of the previous message (“tweet”) in a new message, sometimes with a comment added” (p. 11). Hong et al. (2011)

compare this to the idea of forwarding an e-mail as a form of sharing information or a message with a broader audience. It is also similar to hitting the *share* button on Facebook and other social network venues. In a study about retweeting, Boyd, Golder, and Lotan (2010) argue that there is a direct link between *what* and *why* people retweet. They identified the content of a retweet (the what) to include news they would like to share with others, content they are personally interested in, content they think would appeal to their audience, or content that is meant to call for social action, like protesting, donating, or supporting a cause. Boyd et al. (2010) also listed the reasons people retweet (the why): a) to spread the tweet to a new audience; b) to add a comment to a tweet; c) to show agreement with someone; d) to save the tweet for easier future access; e) to gain more followers; or f) to share the information in the tweet.

One of the characteristics of Twitter is the presence of frequent typos, misspellings, and/or slang (Go, Bhayani, & Huang, 2009). This is encouraged by Twitter's casual nature (Liu, Li, & Guo, 2012) as it is a spontaneous mode of expression for a variety of topics that it is easily accessible through portable devices such as smart phones or tablets. The advent of smart phones increased the ease and speed of access to Twitter as well as other social network sites. In addition, posting unedited tweets makes it easier not to commit to grammatical structure or abide by language norms while typing a tweet.

Conclusion

The ideas presented through the past research explained above forms the context for the present study. I investigated how polyglossic Egyptian users of Twitter replied to existing tweets, and in doing so, which code they used. The aim from doing this was to understand their strategies of taking stance. Code-choice and replying behavior determine the stance which

participants take. I used Du Bois' (2007) framework of analyzing three kinds of stance acts: epistemic, affective, and evaluative.

Stance in the past research has been examined in conversational discourse, in academic writing, and in corpora of both. However, it has not been examined in Discourse 2.0, not on Twitter specifically. In addition, past research focused on blogs. Examining the use of language on micro-blogs is missing, particularly research with a social focus on people and how they express themselves. Further, research examining the online identity of polyglossic Egyptians is scarce. What is even scarcer is research examining polyglossic Egyptian users of Twitter. The only study found examining this is Kossoff (2014). This gap in literature is the reason for undertaking this study.

CHAPTER THREE: METHODOLOGY

Introduction

The current study was an attempt to understanding the relationship between polyglossic Egyptians' online identity on Twitter, especially with regards to stance, accommodation, and the code with which they express their opinions. In this chapter, I will explain the research design, the sample used for this study, the instruments used to address the research questions, and the data analysis techniques used to reach the results.

Research Design

This study is descriptive and exploratory utilizing a qualitative design where I counted frequency to add descriptive data. First, like Kosoff (2014), different codes used by the participants were identified, and their frequency was counted. After that, stance was used as a framework for the qualitative analysis of the data, following the examples of previous studies done by Hyland (2005), Hunston (2007), Johnstone (2007), Jaffe (2009), Bassiouney (2015), and others discussed in the previous chapter. Jaffe (2009) contends that "the introduction of stance into sociolinguistic analysis, especially in conjunction with the field's retheorizing of style, moves the sociolinguistic study of identity into fruitful new directions … the emerging sociolinguistics of stance provides a firm and fertile empirical ground for investigating the linguistic construction of social identity" (p. 41).

Participants

The participant sample comprised undergraduate polyglossic Egyptians in an English-medium university in Egypt. Their ages ranged between 16 and 23 years old, and the number of female participants was almost double that of the males. This university was chosen as the researcher has access to it and its Egyptian students are proficient in both Arabic and English.

Coursework in the university is conducted in English although proficiency levels of the students in Arabic and English vary according to individual differences. For examples, some of them believe their native language to be Arabic while others think it is English. They might describe their own proficiency in either of the languages to be weak, intermediate, or advanced. In all cases, their answers would still count and would be treated equally. There is minimum English language requirement for them to be admitted, so even if they describe their English as weak, it is not really on an objective scale weak.

In this study a DCT is used instead of actual tweets from Twitter because it is not possible to identify a specific population on Twitter. In addition, the demographic information needed for the study will not be available on Twitter and what is available might not be accurate. Furthermore, the earlier explanation of tweets being asynchronous modes of communication is of high importance because bearing in mind that one can take his or her time to think of a tweet before typing it makes the data collected through the DCT to a great extent authentic even though it might not appear on Twitter.

Data Collection Procedures

Access to the participants was achieved by approaching the Student Union (SU) and asking them for assistance conducting the research by forwarding the instrument to their database of students via e-mail. To be able to generalize the findings, the sample that was originally aimed at was to consist of at least 60 students. They were asked to respond to a DCT. Although all students were invited to respond, only the responses of Egyptian students were included in the study as it investigates the use of language by polyglossic Egyptians. The key variable in this study is the frequency of using Twitter. To measure that, various questions are

asked in the demographic section of the DCT given to the participants to fill out to provide relevant demographic data. These questions will be described in the *Instrument* section.

Methods of Data Collection

To collect the data needed for this study, an instrument was used, which was divided into two sections:

1. Demographic Information: This section is a questionnaire where participants answered some questions about themselves and about their Twitter accounts.
2. Tweets: This section comprised the materials where participants were presented with ten tweets and asked to compose a reply to each.

The approval of the Institutional Review Board (IRB) is attached in Appendix C. The consent form which participants signed to indicate their consent to participate in this study is included in Appendix D, and the full DCT is provided in Appendix A.

Demographic information. This section of the instrument is what will help narrow down the choice of the participants. First, there are some questions about the participants' nationality, age, sex, educational background, and proficiency in both English and Arabic languages. While age, gender, and educational background are not deciding variables, questions about them were added in case the data appeared to differ according to either of them. A study about the difference between male and female Twitter users in a corpus of 4.1 million tweets collected from 180,000 users concluded that females used Twitter, tweeted, and were explicit about their gender in the *Bio* section of their profiles more than males (Burger et al., 2011).

Asking about the command of English and Arabic language is important because it could give an explanation as to why polyglossic Egyptians in this study reply in the specific codes they choose. Having full or partial command of either of the languages, yet choosing to express

oneself in only one of them, for example, is a stance taken. Participants who do not have a Linguistic background might not readily think of a distinction between the Arabic codes MSA and ECA. Therefore, the questionnaire (see Appendix A) required knowing what the participants believed their command of the Arabic language to be without further distinction of its types. Had participants asked for more specifications, they would have been advised that their command of ECA was the one intended. As a code, full-command of MSA in Egypt is only required by people who use it in their studies and/or work. It is not a requirement for any kind of communication by those who do not need it in their professional life.

The fact that the population needed for the study is wide and must stay anonymous to the researcher creates a methodological challenge of not knowing anything about the participants. Identifying the stance of a person requires some knowledge of that person. To overcome this challenge, and especially with regards to the topics used in the tweets, two questions were added to the DCT asking some information about the participants' political and religious backgrounds. It is worth-mentioning that this is not a problematic challenge in the study because a person's online identity is different from his or her offline identity, and people often construct different identities in different situation or with different audiences (Paltridge, 2012).

The DCT also includes questions about when participants have joined Twitter, how many followers they have, and how many tweets they have posted. These are important because participants who answer with *a* to question 13 (see Appendix A) will be disregarded. Java, Finn, Song, and Tseng (2007) define a Twitter active user as someone who posts at least one post per week. With websites or apps such as Twitter, a person needs some time to familiarize him or herself with them. This is something that happens by time, and not by the mere fact of creating an account on any medium. For a participant to be eligible for the study, he or she must have

posted more than 100 tweets. The aim is to include users who are familiar with Twitter and have experience posting a good number of tweets. Before developing interest in understanding Twitter and how people use it, I used Twitter rarely and had few posts over a whole year. After developing an interest in Twitter, I began to use it more and posted around 1000 tweets in less than a year.

The reason why the number of followers is not be taken into serious consideration is that on Twitter any user can use certain strategies to get many followers; this can happen over few days and does not mean that a person is a proficient Twitter user. Other people, however, might prefer not to use any of these strategies and have a very small number of followers, yet be frequent Twitter users. A final question asking the Twitter username of the participants is asked in case a closer look at a participant's account is needed later on if the data from his or her replies appear to be interesting. This is an optional question though to allow for a space of freedom if the participants do not wish to share such information.

Tweets. The materials that were used in this study were authentic tweets written by a variety of Twitter users. These tweets were collected randomly from Twitter and the participants were asked to respond to them through a Discourse Completion Task (DCT). When one searches using word(s) or hashtag(s) on Twitter, the results will be all the tweets which have the word(s) or the hashtag(s) written. The process of choosing the tweets which acted as prompts for the participants to type another tweet in reply to each went as follows. First, the tweets were chosen through rather general political, religious, and social topics, like *freedom*, *religious affiliations*, *stress*, *bullying*, *abuse*, *studying*, and *ISIS*. Hundreds of tweets were read then the choice was narrowed down to only ten tweets of which screenshots were taken. These were ones which to a great extent carried a full idea; i.e. can be understood on their own, and which were likely to

arouse debate or strong feelings on their readers. Intensity of emotions is a guarantee for spontaneity in the reply which is needed to make the reply as authentic as possible and, thus, reach a conclusion about how the use of language reflects the identity of its users. Some of them were light in nature. These were intentionally added to give the participants a better feeling of the authenticity of these tweets, since Twitter is in reality full of them. The selected tweets discussed social, religious, and political ideas from different perspectives. None of them held the personal opinion of the researcher who remained objective toward all the presented opinions in the original tweets and the replies of the participants.

The codes chosen for the selected tweets varied. They were written in Arabic, English, Arabizi, Arabized English (AE), and Code-switching (CS) between two or more of these codes. AE is a new code which can be found in the 7th tweet and was presented as an emerging code in Egypt at least, if not in the whole Arab world. This code is the counterpart of Arabizi. It is English written in Arabic letters. The importance of the presented variation of codes is to avoid the participants being affected by the fact that all the original tweets are written in only one of the codes, thus thinking they can only reply in the same one and not another one. Finding all the codes represented in the DCT adds to the validity of the data collected because it frees participants from being limited in the codes they tweet in. Further, it adds to the authenticity of the DCT making it a better representative of a Twitter homepage. See Appendix B for the translation of the tweets written in Arabic or Arabizi. The translation of the Arabic tweets does not always fit into 140 characters because the goal behind translating them is to explain them and not to have them fit as original tweets.

Visual Layout of the tweets. Another factor that can influence the reply of the participants is how authentic the tweets looked. Towards that end, screenshots of the tweets were

taken rather than typing them. This was a means of assuring the participants that these were real tweets which can, in effect, put them more in the feeling of writing a real tweet the way they really did. However, this could result in the replies being affected by the identity of the writer of the original tweet. There are two cases where this influence can happen. First, in case it is a public figure which might put the participants in a stance based on such knowledge. Second, in case the original tweet's writer is a foreigner whom the participants might suspect not to know Arabic, then their reply might be in English just to be understood by the person they are replying to instead of reflecting who they really are. To avoid these possible problems, parts where the name and the profile picture are clear were hidden. Thus, the tweets are to remain anonymous to the participants to avoid the possibility of affecting their reply and which will not to be measured in the study. See the DCT in Appendix A to clarify this idea.

The order of the tweets. The language in which tweets were written was put in mind. They were mixed to avoid having tweets of one code followed by tweets of another code. Second, the ones with the more serious nature were added in the middle separated by less serious ones in order not to cause a feeling of uneasiness among the participants. Furthermore, the tweets which have rather light nature were separated from each other to add the effect of naturalness and spontaneity to the DCT to help it more closely resemble a homepage on Twitter, and thus, make the DCT look as genuine as possible. Starting with an Arabizi tweet was meant to draw the participants' attention that they can reply in any code they prefer, especially in case they did not read the heading of the section carefully, which has clear instructions that: "You are encouraged to reply in any language you normally tweet in, like Arabic, English, a mix of both, or Franco-Arab."

Challenges. Piloting my study rendered valuable insight into few issues which were taken care of during the actual study. I shall explain these issues in case a similar or close study is to be done. Some of these issues were with regards to the questions used in the instrument, others had to do with the chosen original tweets, and some were more general ideas. Following will be a detailed ordered description of each of these issues and/or challenges.

Demographic Information. In the demographic section, the question asking about the participants' nationality has two answers: a) Egyptian and b) Other. Although identifying the *other* nationality is of no importance, since data from non-Egyptians were not going to be used, more interest in volunteers' nationality was given by asking to specify it as a sign of courtesy and respect to their willingness in helping the researcher. Further, although the question asking whether participants had a Twitter account or not seems to be of no use, it was kept in case the DCT in case a participant replies to all the tweets despite not having an account.

Tweets. There were four problems located with the tweets chosen for the pilot study. One of them, for example, had a link for more details about the topic of the tweet. Two of the participants needed to know about the whole argument before replying; hence, tweets which had further links of additional information were not included in the actual study. The second problem was with a tweet which some participants did not understand. Therefore, in the actual study, I tried to avoid tweets that could be problematic. Tweets should not be interpreted; they should be understood by the participants on their own even if the meaning they got is not exactly the one intended. Third, since the DCT was sent online to participants, it was important that the question draws the participants' attention to not exceed 140 characters which is the maximum number of characters one can use in typing a tweet. This is why an additional sentence was added and

underlined under the question of replying to the tweets: “While answering, kindly bear in mind that your replies should NOT exceed 140 characters.”

Data Analysis Techniques

In this section, I will explain in detail the techniques used in analyzing the data collected through the method explained above. First, after the data were collected, the codes of the replies by polyglossic Egytpains on Twitter were identified, and then the frequency of each was counted. Second, the number of codes written for each of the tweets was also counted to find whether respondents accommodated the tweet authors in their replies or not. Third, a close reading of the replies was done to find the strategies of replying among participants identified through repeated patterns. Fourth, Du Bois’ (2007) stance triangle framework was used in the analysis of the participants’ replies to tweets.

Using Du Bois’ framework, I looked at the nature of stancetaking in the participants’ replies. The replies were classified as expressing an epistemic or affective stance. The linguistic elements used to express each of these types of stance were identified qualitatively through providing examples to them while mentioning their significance and explaining how they support stances taken by the participants. These linguistic elements included modality, hedges, adjectives, imperatives, and factual statements. Identifying them adds to the objectivity of the analysis and the results. Finally, and to link all the results reached through the analysis explained above, a relation between code choice, accommodation, and the stance taken by the participants was explained, then linked to polyglossic Egyptians’ Twitter online identity.

CHAPTER FOUR: RESULTS

Introduction

This study aimed at understanding ways in which polyglossic Egyptians take stances through their replies to tweets on Twitter. It answered three main research questions: the first was about the codes used by polyglossic Egyptians in their replies on Twitter, the frequency of use of these codes, and the relation between the choice of code and stance; the second was about the relation between stance and accommodation; and the third was about the replying strategies exhibited on Twitter and the methods through which stance is employed through these replies. To collect the data needed to answer these questions, a DCT consisting of two sections was sent to the participants. The first required providing self-reported data about the participants' Arabic and English proficiency and about their Twitter accounts. The second consisted of screenshots of ten tweets taken from Twitter asking the participants to reply to each of them (see Appendix A for the DCT).

This chapter offers answers to these three research questions with their sub-questions. The first and the second questions which aimed to identify the codes used by the participants and how far they accommodated in their replies were answered through a qualitative approach. In the first question, a frequency count of the different codes participants used to type their replies to the tweets was done, and then the data were explained qualitatively through identifying the relationship between the codes used in the replies and the stance taken by participants. The second question was answered by measuring the degree of accommodation with which respondents replied and relating accommodation to the stances they took. Qualitative analysis also was done for the third question which was answered through identifying the patterns of replying and the stance-markers used for these replies.

Filtering the Replies

The instrument (the DCT) was sent to a huge database of Twitter users who were students at an English-medium university in Egypt. The responses received were 109 from the database of university students. The responses from non-Egyptians were filtered since the study does not measure the use of non-Egyptians on Twitter. Those who responded as having only posted between 0 and 100 tweets were filtered too, since, as clarified previously in the third chapter, there is a difference between a passive and an active user of Twitter, and this study only targeted active users. Thus, the number of responses used for the analysis was 81. In other words, 810 tweet replies from the 81 participants comprised the saw data collected for this study.

First Research Question, Sub-Questions a & b: Codes Used in the Replies

The second step after the data were collected was to identify the codes of the 810 replies by the participants. The codes identified were: Arabic, English, Arabizi, Arabized English (AE), code-switching (CS), another language, "sounds, emoticons, and/or emojis," (SEE) and no reply. Table 4.1 shows these codes, the number of replies written using each of them, and its percentage of use among the 810 replies received.

Table 4.1

Codes Used by Participants in their Replies

Code Used	Count (%)
Arabic	74 (9.1%)
English	414 (51.1%)
Arabizi	170 (21%)
AE	3 (0.4%)
CS	60 (7.4%)
Another language	2 (0.2%)
SEE	72 (8.9%)
No reply	15 (1.9%)

For an understanding of AE, CS, and SEE, see the "Definition of Constructs" section in chapter one. When a participant replied in one of the first four codes (Arabic, English, Arabizi, or AE) along with a sound, an emoticon, and/or an emoji, the reply was coded as belonging under the text written. In other words, if a reply, for example, was written in English along with an emoticon, it was coded as an English reply. The code SEE was treated as supplementary feelings added to enrich the text, and hence was given secondary importance. It was given to replies written only in one or more of them. For example, when a respondent replied using an emoticon or using a sound and an emoji, the reply was coded as SEE. The code *no reply* was given to replies with several dots or dashes indicating that the respondent did not wish to reply but had to fill out the blank space of the reply so as to be able to submit the DCT. *Another language* was added as a code since two of the replies made by two different participants were in French.

As shown in Table 4.1, the first research question was answered by offering the order of frequency of the codes used by the participants as follow: English by 51.1%, Arabizi by 21%, Arabic by 9.1%, SEE by 8.9%, CS by 7.4%, No reply by 1.9%, AE by 0.4%, and Another language by 0.2%. The fact that almost half of the replies were made in English indicates that respondents, who earlier identified themselves as Egyptians through their nationality, do not believe that using a code other than Arabic precludes them from being Egyptian. Their online identities on Twitter were shaped more by English as a code than it was by Arabizi or Arabic (the following two most commonly occurring codes). Replying in Arabizi which occurred second to English in frequency is not surprising given that the age group of the participants ranged mostly between 16 and 23 years old. Replying in SEE by 8.9% indicates that their online identities were also often shaped by one of these sound or visual devices as a marker of stance, a

stance of not wishing to communicate with language if the same message can be communicated in a visual or a sound. Choosing not to reply is a conscious act of not wanting to. Hence, it can be considered a stance of expressing silent opinion, of non-approving, or of ignoring what someone has posted.

First Research Question, Sub-Question c: Political and Religious Identification

Through this section, the last sub-question of the first research question was answered. Two of the demographic questions in the demographic section of the DCT (see Appendix A) required the participants to choose how they would identify themselves politically and religiously. The aim of these questions was to see if there was a relation between self-identification and the choice of code, and hence how this affected the stance taken by the participants. See Table 4.2 for the use of codes in relation to participants' political identification.

Table 4.2

Political Identification of Participants

Code Used	Liberal	Conservative	None of the above
Arabic	50 (14.3%)	1 (2.5%)	23 (5.5%)
English	178 (51%)	21 (52.5%)	215 (51.2%)
Arabizi	66 (18.8%)	14 (35%)	90 (21.4%)
AE	2 (0.6%)	0 (0%)	1 (0.2%)
CS	25 (7.1%)	2 (5%)	33 (7.9%)
Another Language	1 (0.3%)	0 (0%)	1 (0.2%)
SEE	25 (7.1%)	2 (5%)	45 (10.7%)
No reply	3 (0.8%)	0 (0%)	12 (2.8%)
Total Participants	35 (43.2%)	4 (5%)	42 (51.8%)

As shown in Table 4.2, using English as a code for replying got almost the same percentage (51-52%) whether the participant identified himself or herself with being liberal, conservative, or neither. Comparing the use of Arabic to the use of Arabizi, Arabic was used more (14.3%) by participants who identified themselves as liberals, while Arabizi was used more

(35%) by participants who identified themselves as conservatives. The data show that there is not a clear correlation between any one code and tweet author's political identity. This suggests that polyglossic Egyptians' code of choice does not reflect their political identities.

Table 4.3 shows the results of participants' religious self-identification and the code use.

Table 4.3

Religious Identification of Participants

Code Used	Religious and liberal	Religious	Non-religious	None of the above
Arabic	46 (9.8%)	17 (8.5%)	5 (7.1%)	6 (8.6%)
English	234 (49.8%)	107 (53.5%)	35 (50%)	38 (54.3%)
Arabizi	106 (22.5%)	36 (18%)	16 (22.9%)	12 (17.1%)
AE	2 (0.4%)	1 (0.5%)	0 (0%)	0 (0%)
CS	31 (6.6%)	17 (8.5%)	3 (4.3%)	9 (12.6%)
Another Language	2 (0.4%)	0 (0%)	0 (0%)	0 (0%)
SEE	44 (9.4%)	14 (7%)	9 (12.9%)	5 (7.1%)
No reply	5 (1.1%)	8 (4%)	2 (2.8%)	0 (0%)
Total Participants	47 (58%)	20 (24.7%)	7 (8.65%)	7 (8.65%)

The use of different codes according to self-reported religiosity, as shown in Table 4.3, occurred almost in similar distribution. Across the four religious identifications, English was the code most used by all participants ranging from 49.8% to 54.3% of the replies. Use of Arabic ranged between 7.1 to 9.8% of the replies, and the use of Arabizi ranged from 17.1% to 22.5% of the replies. Thus, this data suggests that a relationship between religiosity and the choice of code does not directly exist.

Second Research Question: Degree of Accommodation

To answer the second research question, the codes of the replies to each of the tweets were quantified followed by a qualitative analysis of these numbers. In the context of the present study, to accommodate means to use the same code which the Twitter user wrote his or her tweet with. Tables 4.5, 4.6, 4.7, and 4.8 present the tweet numbers—as they were used in the DCT (see

Appendix A) and as shown in Table 4.4—the code each was written in, and the number and percentage with which each of the codes of the replies was written. See Appendix B for a translation of the Arabic and Arabizi tweets.

Table 4.4

Tweets Numbers and the Codes of each

Tweet #	Tweet	Tweet Code
1	“3yz akalemek f hewar” BETWA2A3 ALBAK WALAHY!	Arabizi
2	More people are concerned with why women stay I abusive relationships than why men are abusing women	English
3	#داعشي_حتي_النخاع داعش لقب اطلقوه الحاقدين الصحيح يسمي الدولة الاسلامية و مواطنها يسمي (مسلم) و ليس داعشي	Arabic
4	The continuous struggle between “I know me better” and "اكيد مش كل الناس غلط و انا اللي صح"	CS (English & Arabic)
5	Looking at the old conversations w b2ol ana ezay kont keda ? Eh da bgd 😂	CS (English & Arabizi)
6	Your religion might be holding back your development as a person. #atheist #atheism #freethinker #humanism	English
7	البروكراستينيشن صابتني و تأجيل الديدلاين نجّاني #مثل_صيني	CS (AE & Arabic)
8	الفرق بين اللي هتلر عمله و اللي ترامب نفسه يعمله، اننا مش هنعرف نعمل الcomeback الفشيخ بتاع اليهود و نحكم العالم بعدها	CS (Arabic & English)
9	The only thing that can set everything straight in Egypt is a proper LGBT + Sex work movement	English
10	#سبب_عنوسه_البنات "بكمل دراستي"	Arabic

Table 4.5

Numbers & Percentages of Code Replies to Tweets 1, 2, & 3

Tweet#/Code Used	T1/Arabizi	T2/English	T3/Arabic
Arabic	3 (3.7%)	0 (0%)	22 (27.2%)
English	25 (30.9%)	73 (90.1%)	32 (39.5%)

Arabizi	42 (51.9%)	3 (3.7%)	19 (23.4%)
AE	0 (0%)	0 (0%)	1 (1.2%)
CS	7 (8.6%)	4 (5%)	3 (3.7%)
Another Language	0 (0%)	0 (0%)	0 (0%)
SEE	4 (4.9%)	0 (0%)	2 (2.5%)
No reply	0 (0%)	1 (1.2%)	2 (2.5%)

As shown in Table 4.4, the first tweet was written in Arabizi. It received 51.9% of the replies to it in Arabizi indicating a choice of accommodation by more than half of the respondents to the Twitter author. The second tweet received 90.1% of the replies in the same code of the tweet (English) indicating a clear accommodation by the respondents to the Twitter author. The third tweet was written in Arabic and was responded to by 27.2% in Arabic, 39.5% in English, and 23.4% in Arabizi. Hence, there was less accommodation to the tweet writer.

Table 4.6

Numbers & Percentages of Code Replies to Tweets 4 & 5

Tweet#/Code Used	T4/CS (Arabic & English)	T5/CS (English & Arabizi)
Arabic	7 (8.6%)	6 (7.4%)
English	49 (60.5%)	31 (38.3%)
Arabizi	16 (19.8%)	25 (30.8%)
AE	0 (0%)	0 (0%)
CS	5 (6.2%)	8 (9.9%)
Another Language	0 (0%)	0 (0%)
SEE	3 (3.7%)	11 (13.6%)
No reply	1 (1.2%)	0 (0%)

The fourth tweet was written in CS between Arabic and English to which 6.2% only of the respondents code-switched. The rest wrote in only one of these codes indicating a lower degree of accommodation. The fifth tweet was written in CS between English and Arabizi to which 9.9% accommodated by code-switching, while the rest wrote in only one of the codes.

Table 4.7

Numbers & Percentages of Code Replies to Tweets 6 & 7

Tweet#/Code Used	T6/English	T7/CS (Arabic & AE)
Arabic	2 (2.5%)	11 (13.6%)
English	69 (85.2%)	18 (22.2%)
Arabizi	3 (3.7)	20 (24.7%)
AE	0 (0%)	0 (0%)
CS	3 (3.7%)	4 (5%)
Another Language	1 (1.2%)	0 (0%)
SEE	0 (0%)	25 (30.8%)
No reply	3 (3.7%)	3 (3.7%)

The sixth tweet was written in English, and it received 85.2% of the replies in English indicating a clear degree of accommodation to the tweet writer. The seventh tweet was written in CS, and it received only 5% of the replies in code-switching. The rest of the respondents used only one code to reply to it.

Table 4.8

Numbers & Percentages of Code Replies to Tweets 8, 9, & 10

Tweet#/Code Used	T8/CS (Arabic & English)	T9/English	T10/Arabic
Arabic	9 (11.1%)	1 (1.2%)	13 (16%)
English	33 (40.8%)	62 (76.6%)	22 (27.2%)
Arabizi	13 (16%)	7 (8.6%)	22 (27.2%)
AE	0 (0%)	1 (1.2%)	1 (1.2%)
CS	4 (4.9%)	5 (6.2%)	17 (21%)
Another Language	0 (0%)	0 (0%)	1 (1.2%)
SEE	19 (23.5%)	3 (3.7%)	5 (6.2%)
No reply	3 (3.7%)	2 (2.5%)	0 (0%)

Like the seventh tweet, the eighth was also written in CS, and it received 5% of accommodated replies in CS. Like the previous tweets which were written in English, tweet number nine received also a high percentage (76.6%) of replies in the same code (English).

Tweet number 10 which was written in Arabic received 16% of the replies in Arabic, 27.2% in English, 27.2% also in Arabizi and 21% in CS.

Third Research Question: Replying and Stance-taking Strategies

Answering the third research question was done through a qualitative data analysis of six of the ten tweets in the DCT. These were the second, the third, the sixth, the eighth, the ninth and the tenth tweets (See Appendix A). Choice of these tweets specifically was done because these were the more provocative tweets which evoked more feelings on the part of the participants. These feelings were expressed through longer replies and, hence, more room to conduct linguistic analysis.

Through the replies to the screenshots of the tweets in the DCT, it was found that respondents had three shifts of focus upon which they based their replies: the self, the addressee, and the topic. Following will be a detailed explanation for these three points of focus and for the acts and markers of stance found within each.

Self-based replies. Some respondents chose to mention themselves as they replied to the tweet marking a stance of very clear recognition of oneself and one's ideas as the main focus and the main point in the reply. The stancetakers' subjectivity was invoked through this act of positioning through linguistic elements including: a) the pronouns *I*, *me*, and *my*; b) equivalent Arabic pronouns, like أنا (literally meaning "I") and the first person possessive suffix ي (literally meaning "to me"); c) English and Arabic verbs with the pronouns added or omitted which refer back to oneself; and d) English adjectives also with the pronouns added or omitted which refer back to oneself. Consider these replies to the tweets as examples. The subjectivity elements explained are highlighted in yellow for easy recognition.

Second tweet.

(1) Idk if that's a fact or not but I think people should be more concerned about a solution not statistics

(2) mmm never thought of that

(3) totally agree

Third tweet.

(4) msh fahma
miʃ fahma
do-not understand-I-f.
I don't get this

(5) اتمنى أن يتعلم المرء ما يمثل الاسلام بدلاً من إتباع المتأسلمين لأغراض سياسية مثل دعش
ʔatamanna ʔan yataʕallam ʔal marʔ ma yumaθel ʔal Islam badalan min
Hope-I that learn the person what represent-s the Islam instead of

ʔitiba:ʕ ʔal mutaʔslimi:n li ʔaġra:d siyasiya miθl daʕiʃ
following the Islamists-pl for reasons-pl political like Daesh

I wish one would learn what Islam is rather than follow so-called muslims with political agendas like Daesh/ISIS

Sixth tweet.

(6) My religion is my only reason and motivation to my development

Eighth tweet.

(7) ترامب ده مش مسببلي اي قلق خالص
Trump dah miʃ misabibli: ʔy ʔalaʔ xa:lis
Trump-m this not cause-me any worry at all
Trump doesn't worry me at all

The stancetakers (the respondents) in the self-based replies chose to mention themselves to express their stances. The first reply starts with *I* written through the acronym *Idk* (for "I don't know"). The second and third replies show examples of English verbs whose pronouns—referring back to the self—are omitted, like *thought* and *agree*. The fourth is similar to the

second and third examples, but the verb is an Arabic one written in Arabizi also with the pronoun omitted. The first word in the fifth reply is an Arabic verb which means *I hope*. The sixth reply had three "my" pronouns where the respondent referred back to him or herself and opinion. The seventh reply had an example of the Arabic firxt possessive suffix "ي" added at the end of the verb which, as explained above, literally means "to me."

The third example reflects a stance of agreement and alignment with the tweet writer marked through the verb *agree* and the booster *totally*. The rest are examples of epistemic stances by the stancetakers. In examples 1, 2, and 4, the epistemic stance is marked through the use of the verbs which mark a degree of knowledge or understanding, like *don't know* (in *Idk*), *think, thought,* and *fahma* (meaning "I understand"). However, in examples 5, 6, and 7, it was marked through stating an opinion as if fact. In other words, the stancetakers tweeted their ideas in the form of stating facts.

Addressee-based replies. Some respondents chose to respond to the original tweet writers directly through referring to or addressing them, asking them questions, giving them advice, and/or questioning their ideas or opinions. To do that, respondents used linguistic elements like: a) the imperative tense of verbs directed to the tweet writer; b) direct references using *you* or its Arabic equivalent ك, the second person possessive suffix (meaning "your"); and c) question marks which denote asking the tweet authors question(s) about the content of what they had delivered through their original tweets. The following are examples taken from the participants' data to clarify this shift of focus. Elements of addressee-based focus are highlighted in green.

Second tweet.

(8) Don't generalize

(9) By people you mean who?

Third tweet.

(10) erhaby we motetrf
ʔirhabi wa mutatarrif
terrorist and extremist
What a terrorist and an extremist

(11) Daesh = dawlaislameya? What kind of joke is this?
Daesh = dawla ʔislamiya? What kind of joke is this?
Daesh = state-f. Islamic ? What kind of joke is this?
Daesh = an Islamic state? What kind of joke is this?

Sixth tweet.

(12) Don't hashtag atheist and freethinker bas wenaby habeeby.
Don't hashtag atheist and freethinker bas wenaby ḥabi:bi .
Don't hashtag atheist and freethinker just for the love of the prophet my love.
For God's sake don't hashtag atheist and freethinker, darling.

Ninth Tweet.

(13) You really think so?

(14) Sex work as in prostitution?

Tenth tweet.

(15) سسب عنوسة البنات كثرة الناس الي بيفكروا زيك و بيلقبوا الناس على مزاجهم
Sabab ʕunu:sit ʔil banat kaθrat ʔil nas ʔilli beyfakaru: zayak
Cause spinsterhood the girls increase the people who think like-you-m.
wi biylakabu ʔil nas ʕala mazaghum
and label-pl. the people on

Spinsterhood is caused by people who think like you and label others as they please

(16) Wi yaretik fal7a feeha
Wi ya ritik falḥa fi:ha
And hopefully-f successful-you-f at
And you're not even a good student

All the addressee-based tweets were epistemic stances expressed through sarcastic comments, offering advice, questioning ideas, or stating facts. In reply number 10, direct

evaluation is invoked where the tweet writer is treated by the stancetaker as the object of stance who is evaluated through the use of the adjectives: *terrorist* and *extremist,* which are the stance markers in this reply. This evaluation is accompanied by a clear stance of disalignment with the tweet writer.

Replies numbers 8 and 12 have the imperative form of the verb "do" ordering, or rather advising the tweet writers with what they should not do, which marks a stance of superiority coming from better understanding and more wisdom in making better decisions. Another stance marker in example 12 is the sarcastic use of a term of endearment *habeeby* which literally means "my love" to address a stranger. This also supports the same stance of superiority with a patronizing tone. In examples 9 and 13 the pronoun "you" was used to address the tweet writers directly. In the same two examples, along with examples 11 and 14, question marks were used to indicate asking the tweet authors questions. Questions 9 and 14 were examples of apparently real questions requesting additional information from the authors of the tweets. However, examples 11 and 13 carry clear extra meanings of sarcasm, of questioning the ideas of the tweet authors marking disalignment with them. The word *joke* in example 11 supports this idea.

In examples 15 and 16, the Arabic second person possessive suffix ك (meaning "your") was added to the end of words as a means of addressing the tweets writers. Using an epistemic stance in example 15 where the reply carries a clear piece of information which the participant sounds very confident of and which is different from that in the tweet, the stancetaker claims the stance of higher degree of knowledge than that of the tweet writer. In the last example (example 16), the marker of stance is sarcasm, and it is indicated through the verb "fal7a" (literally meaning "successful") which has a direct accusation to the tweet writer who is assumed to be a female in this tweet of being not just a spinster but an unsuccessful person too.

Topic-based replies. Some respondents chose to directly respond to the topic of the tweet, rather than referring to themselves or addressing the tweet author. They referred to and expressed their opinion about the topic itself. In most cases, this was a way of disaligning with the opinion expressed in the tweet they were replying to. Methods of giving a topic-based reply were: a) starting the reply with a logical connector as a way of completing the tweet; b) responding with an elliptical reply; and c) responding with a stand-alone reply. These methods will be explained and given examples for below.

Completing the tweet. Some participants started their replies as if they were completing the tweet itself. The linguistic elements for this were starting the reply with logical connectors like *and*, *because*, *cause*, *unless*, and *3ashan* (meaning "because"). Examples 17, 18, 19, 20, 21, and 23 show these elements which are highlighted in light blue for easier reference. Another marker was starting the reply with an additional piece of information to complete the tweet. Example 22 is an example of this marker where the respondent started his or her reply by giving an additional piece of information to complete the one in the tweet. In that reply, the respondent was explaining another difference ("Elfar2 eltany") between Trump and Hitler from his or her point of view.

Second tweet.

(17) Because we're better in moving away from the problem rather than solving it. #Fucklogic

(18) 3shan it's easier to influence and convince the abused than convince the abuser, and ultimately make a change .
ʕaʃan it's easier ...
Because it's easier …
Cause it's easier …

(19) And some other people don't consider why women abuse men

(20) because it's always up to the girl and what she does not what the guy does

Sixth tweet.

(21) Unless that religion precisely encourages development in order to benefit society. Islam for you (:

Eighth tweet.

(22) Elfar2 eltany was that Hitler was actually a charismatic and strong leader, Trump is an idiot supported by baffoons...
ʔil far? *ʔil tani* *was that …*
The difference-m the second-m was that …
The other difference was that …

Ninth tweet.

(23) And sex education

In all the replies above, the stance expressed was an epistemic stance of using the simple present or past to express an opinion. They were all presented as if factual statements by the stancetakers along a scale of agreement (as in examples 17, 18, 19, 20, 22, and 23) or disagreement (as in example 21). The epistemic stance in example 22 was reinforced with a stance of agreement and alignment with the tweet writer through approving of the information presented in the tweet. In addition, direct evaluation was expressed in the same example through the use of the adjectives *charismatic*, *strong*, and *idiot* as its markers.

Responding with an elliptical reply. The second type of topic-based reply was achieved by commenting on the topic of the tweet in a dependent reply. In this type of reply, a general judgment on the tweet's content was expressed. It is described as elliptical because it cannot be understood without the context of the original tweet. To clarify, it cannot semantically be understood or stand alone without the original tweet to which it is a reply. This type can be very short in length, and examples include *true, good point, yes* to agree with the tweet, *no, never* to disagree with it, *no reply, N/A, no comment* to abstain from replying, or other comments which

express an indirect agreement or disagreement through a judgment on the content. Examples are presented as follow.

Second tweet.

(24) Deep people everywhere

(25) Nas gahla
Nas gahla
People ignorant
Ignorant people

(26) Amen!

Third tweet.

(27) bala araf
bala ʔaraf
no disgust
disgusting

(28) Pathetic slogan

(29) قشطة
ʔiʃta
cream
cool

(30) شايفين المخدرات ممكن تعمل ايه؟
Shayfi:n ʔil muxdara:t mumkin tiʕmil ʔih?
See-you-pl the drugs-f can do what?
See what drugs can do to you?

Sixth tweet.

(31) #تنميه_الحاديه
Tanmiya ʔiḥaldiya
Development atheistic
#Atheist_Development

(32) Bullshit

(33) irrelevant

Eighth tweet.

(34) Tweet of the year

(35) Yareitt
ya rit
hopefully
Hope so

(36) Dih 7a2i2a!
Di: ḥaʔi:ʔa
This-f truth-f
True

(37) #estoryarab
ʔustur ya rabb
cover-m God
#LordHaveMercy

In the above examples, the stancetakers used epistemic stance occasionally supported by direct evaluation and thereby aligned or disaligned themselves with the tweet authors. Evaluating the objects of stance was done using adjectives as stance markers, like *deep*, *gahla* (meaning "ignorant"), *pathetic*, *قشطة* (meaning "cool"), *bullshit*, and *irrelevant*, and the noun *araf* (meaning "disgust"). In examples 31, 34, 35, 36, and 37, epistemic stance was expressed through presenting an opinion as if it were a fact or hoping for something. An evaluative act was expressed through an epistemic stance in example 30 where a total disalignment with the original tweet was achieved by sarcastically accusing its author of being on drugs. Another direct evaluation occurred through an epistemic stance in the example which read "Tweet of the year" where the respondent aligned with the tweet author by endorsing the tweet and agreeing with its content.

Responding with an elliptical reply can also be rather long where a new opinion is formed and expressed in words while referring to the original tweet. The test to identifying that this is an elliptical reply is that it cannot stand alone; it needs the context of the tweet to be

understood. In the following examples, the use of demonstratives, like *both*, *this*, ده(meaning "this"), and *that,* and the use of pronouns, like احنا (meaning "us") and *it* are what make these replies elliptical ones. They are semantically dependent in the sense that they need the original tweet to be understood.

Second tweet.

(38) people should be concerned with both to find a solution

(39) Men should be educated regarding this matter

Third tweet.

(40) This, right here, is everything that is wrong with this world.

Eighth tweet.

(41) ma ymkn net8ayar ba3d ma y7otona fel m7ra2a ☺
ma yimkin nitɡayyar baʕd ma yḥotina fil maḥraʔa ☺
no possibly change-we after no put-us in holocaust☺
Maybe we'll change after we're thrown into the holocaust ☺

(42) This is wrong & shocking on so many levels!

(43) ده لو بعقلية ترامب العالم مولعش اصلاً.
Dah law biʕakliyit Trump ʔil ʕalam mawliʕʃ ʔaslan.
That if brains Trump the world set in fire in the first place.
What world? It'll be already burned with a mindset like Trump's.

(44) احنا حاكمين بلد واحدة بالعافية..عالم ايه بس
ʔiḥna ḥakmi:n balad waḥda bil ʕafya .. ʕalam ʔih bas
We rule-pl country one with difficulty.. world what just
We can barely rule over one country let alone the whole world

Ninth tweet.

(45) it's not our main problem awareness should be raised first

(46) that is definitely not the thing that will set egypt "straight", weve got bigger problems

(47) there are way more important things than this

Tenth tweet.

(48) it depends on the girls perspective msh aktr
it depends on the girls perspective miʃ ʔaktar
it depends on the girls perspective not more
it only depends on the girls' perspective

(49) ده مش سبب يخليها عانس بالعكس دي حاجة بترفع من مستواها في المجتمع
Dah miʃ sabab yixali:ha ʕanis bil ʕaks di ḥaga bitirfaʕ min
This not reason make-her spinster on the contrary this something raises from
mustawaha fil mugtamaʕ
status-her in society
That doesn't make her a spinster. It elevates her social status

In all the examples above, an epistemic stance is taken through the use of the simple present and the modal *should* where the stancetakers again claimed the position of knowledge and superiority by giving advice on what should and should not be done (through the modal *should*) and by stating their opinion as if it were a fact (through the simple present). In examples 40 and 42, this epistemic stance was reinforced by a direct evaluative act through the use of adjectives as stance-markers, like *wrong* and *shocking*. In example 46, the epistemic stance was reinforced through the use of the booster *definitely*. The Arabic replies, numbers 43 and 44 used sarcasm as a stance marker where the respondents showed disalignment with the tweet writers through making sarcastic comments about the content of their tweets. Example 41 showed the use of a hedge (*ymkn* meaning "might") which is also an epistemic stance but one of less assertiveness where the stancetaker did not claim absolute certainty of the information he or she is presenting.

Responding with a stand-alone reply. This was done when the reply had no reference to the tweet at all. In other words, the reply seemed like an independent tweet, as if the respondent were not replying and was just expressing an opinion about the topic in a fresh tweet. To measure whether a reply fell in this category, the test was to think of it as a stand-alone tweet. If

it proved to be complete and meaningful standing alone, the tweet fell in this category. Examples of this tweet type follow.

Second tweet.

(50) If a woman is strong enough, she won't tolerate being degraded/abused.

(51) Men abusing women is inexcusable, the PHYSICAL harm has to be stopped first before going in detail about the psychological one.

(52) "Like a compass needle pointing North, man's accusing finger always finds a woman. Always." (:

Third tweet.

(53) Da3sh are not muslims and muslims do not represent da3sh.
Daʕiʃ are not muslims and muslims do not represent daʕiʃ.
Daesh are not muslims and muslims do not represent daesh.
Daesh are not muslims and muslims do not represent daesh.

(54) Isis dol kol el bo3d 3an el Islam. Dol erhabeyeen.
ISIS dul kul ʔil buʕd ʕan ʔil ʔisla:m. Dul ʔirahbiyi:n.
ISIS these all the distance from the Islam. These terrorists.
ISIS are the farthest thing from Islam. They are terrorists.

Sixth tweet.

(55) الدين موجود عشان يرتقي بالناس و الاله الذي بقلل من انسانيه و يحد من افكار و ابداع اللي بيعبدوه و لا يستحق ان يعبد و العبادة علاقه حب #الله_الذي_احبه
ʔil di:n mawgu:d ʕa ʃan yartaqi bil nas wi ʔil ʔilah allaði yuqallil
The religion present because rise-up with people and the god who degrades-m
min ʔinsani:h wa yaḥid min ʔafkar wa ʔibdaʕ illi: byaʕbudu:h
from his-people and limits-m from thoughts and creativity who-pl worship-him-pl
wa la yastaḥik ʔan yuʕbad wa ʔal ʕibada ʕilakit ḥub
and not deserve-m to be-worshipped-m and the worship relationship love
#ʔalla:h_ʔallaði_ʔuḥibuh
#Allah _who _I-love-him

Religion exists to help people rise. A god who dehumanizes and limits the ideas and creativity of his worshipers isn't worthy of worship. Worship is love.
#The_Allah_I_Love

(56) Any religion aims to make humans better, so it is not religion that holds the person back

Eighth tweet.

(57) Hitler w trump men aswa2 el ash5as el etwaladet 3ala wagh el ard.
Hitler wi Trump min ʔaswaʔ ʔil ʔa ʃa:s ʔil ʔitwaladit ʕala wagh ʔil ʔard
Hitler and Trump from worst the people who were-born on face-n the earth
Hitler and Trump are two of the worst people ever born on the face of earth

(58) trump is a dumbass and will never win

Tenth tweet.

(59) Sabab el fa2r wel takhalof wel gender discrimnation wel men dominated society: "Mesh lazem akammel derasty"
Sabab ʔil faʔr wi ʔil taxaluf wi ʔil gender discrimination
Reason the difference and the backwardness and the gender discrimination
wi ʔil men dominated society: "Miʃ lazim ʔakamil dirasti"
and the men dominated society: "Not necessary I-complete education-my"

The reason behind poverty, backwardness, gender discrimination, and men dominated society: "I don't have to finish my education"

(60) كلمة عنوسة دي كلمة متخلفة
Kilmit ʕunu:sa di kilma mutaxalifa
Word spinsterhood this-f word-f backward-f
The word "spinsterhood" is very backward

(61) سبب تخلف المجتمع :هدف حياة البنت إنها تتجوز.
Sabab *taxaluf ʔil mugtamaʕ : hadaf ḥayat ʔil bint ʔinnaha titgawiz.*
Reason backwardness the society : goal life the girl that-f gets-married-f.
The cause of society's backwardness: Girls' life purpose is to get married.

Like previous replies, epistemic stance was detected through the following assertive stance markers: the simple present and future tenses, the use of modals like *has to*, the use of adverbs like *always*, and the use of negation like *not* and *never*. Occasionally, the epistemic stance was supported by direct evaluation expressed through the use of adjectives like *erhabeyeen* (meaning "terrorists"), *dumbass*, and *متخلفة* (meaning "backward"), and the use of comparatives like *better* and *aswa2* (meaning "worse"). The evaluative act was also expressed in

an assertive way. In all these examples, stancetakers presented their ideas in a fact-like way marking an epistemic stance of knowing and a position of superiority.

Two-part replies. In some cases, although infrequently, two or more focal points occurred within the same tweet, consisting of two or more parts and divided by full stop(s), question mark, or comma. In the following examples, the self-based parts are highlighted in yellow, the addressee-based parts are highlighted in green, and the topic-based parts are highlighted in light blue. Linguistic elements which identified each of them are like those explained above for each of the points of focus.

Second tweet.

(62) I agree. We need to get to the root of the problem first.

(63) hmm.. not quit sure. there's always a reason.

Third tweet.

(64) ايه الاسلام في انك تقتل و تدبح ....داعش ابعد ما تكون عن الاسلام
ʔih ʔil ʔislam fi: ʔinnak tiʔtil wi tidbaḥ .... Daʕiʃ ʔabʕad ma taku:n
What the Islam in that-you-m kill-m and slaughter .... Daesh farthest not be-f
ʕan ʔislam
from Islam

There is nothing Islamic about murder and slaughter. Daesh has nothing to do with Islam.

(65) Daesh are not Muslims, I'm Muslim

(66) المسلم من سلم المسلمون لسانه و يده. و حضرتكوا ايديكوا طايله كله ماشاء الله
ʔal muslim man salam ʔal muslimu:n min lisanuh wa yaduh.
The Muslim-m who avoid the Muslims-m from tongue-his-m and hand-his-m.
Wi ḥadaratku ʔidiku tayla kulu ma ʃaʔ ʔallah
And you-pl hand-your-pl reaches-f all not want-m Allah

A true Muslim is one from whose tongue and hand, all other Muslims are safe. No one is safe from your hands.

Sixth tweet.

(67) highly possible .. but have you considered lack of it could also negatively affect you

(68) NEVER. my religion makes me develop better as a person #ISLAM

Ninth tweet.

(69) Set Egypt straight? well, thats funny

(70) What do you mean by sex work movement? Do you mean sex ed? I'd totally agree if that's what you meant.

Tenth tweet.

(71) So untrue. I personally know a lot of women who are educated or still getting education and are married.

(72) Education comes first in my life. Marriage is not as important as having a good education 5ales.
Education … xalis.
Education … at all.

Education comes first in my life. Marriage is not nearly as important as having a good education.

In the above examples, stancetakers expressed an epistemic stance by the use of stance markers, like the simple present and the use of the modal *need to* in example 62. This stance, like the ones before, was an assertive one where the stancetakers claimed authority through stating their opinions in fact-like ways. The only two instances of unassertive epistemic stance can be seen in examples 63 and 67 through the use of hedges (*not quite sure*, *possible,* and *could*) to express lack of certainty. The epistemic stance in example 62 was reinforced by an agreement stance where the stancetaker aligned with the tweet writer using the verb *agree*. In examples 69 and 71, the adjectives *funny* and *untrue* were the markers of an act of evaluation whereby the stancetakers evaluated the objects of stance negatively to express their disalignment with the tweet writers.

CHAPTER FIVE: DISCUSSION

Introduction

Scientific advances have led to the development of fields of study, like Sociolinguistics and Discourse Analysis. Twitter and other social media platforms are considered very fertile area for research. They have provided rich data that has contributed to the notions of identity and code. There is an established relationship between both notions, one that can be analyzed and understood in multiple ways. Online media of communication constitute some of these ways where the code we use is considered a means of building an online identity which can be very different from that we have established offline.

This study explored the identity of polyglossic Egyptians on Twitter constructed through their code-choice and the strategies they use to take stance in their tweet replies. Three research questions were posed, the answers of which explained the codes used by polyglossic Egyptians on Twitter, the relationship between the code use and the users' stance, accommodation as an act of stance exhibited through the code use, strategies for replying, and the methods of taking stance in these replies. This final chapter will briefly explain the results reached through answering these research questions then focus on discussing these results in detail.

Discussion of the Results

According to Hunston (2007), stance can be best analyzed when we are able to quantify some results along with providing some qualitative data. Together they provide us with the right context which helps us understand stance better. This was the method used for this study and below is the discussion of the results in relation to stance.

Code-choice and stance. Stance is defined by Bassiouney (2014) as how one positions himself or herself in relation to others, situations, or things. This stance can be achieved through

code choice. According to Jaffe (2007), code-choice is an act of taking stance. Thus, the stancetakers' code choice was examined in the study. The most dominant code among the 81 participants was English (51.1%), followed by Arabizi (21%), Arabic (9.1%), SEE (8.9%), CS (7.4%), AE (0.4%), and another language (0.2%).

From these data, it can be suggested that the identity of the participants is hugely affected by English as a code and a linguistic resource. Some of them even replied to all the tweets using only English. This could be the effect of the environment in which they live daily (being students in an English-medium university), but it can also be interpreted as being different. This difference can be considered a stance of disalignment with the average Egyptian who does not master English. In Egypt, differences in the type of education people receive leads to differences in social class, and can even determine whether they will be able to communicate and understand each other or not (see Bassiouney, 2009).

Further, Arabic and Arabizi are basically the same code but with different orthography, either Arabic characters or Latin characters. The fact that Arabizi replies were almost double the Arabic replies supports literature which looked into the reasons for preferring Arabizi to Arabic. Whether it is because of the age group this study examined, or because they think it is closer to everyday language, has no rules or correct grammar, easier in typing, looks cooler, or conveys emotions and feelings better through upper and lower cases, participants preferred replying to tweets in Arabizi to replying to them in Arabic.

Political and religious identification and stance. Because part of the construction of a person's identity is his or her ideology, it was important to link code-choice to the stancetakers' political and religious ideologies and to their identities. Findings from the results of code use in relation to participants' political and religious self-identification stand in contrast to assumptions

in the media and some literature like Suleiman's (2003)[4] that conservatives use Arabic more and foreign language less than liberals. At least on Twitter, data from this study show that polyglossic Egyptians use English the majority of the time regardless of their political self-classification. Although political identification proved to make no difference in the percentage of use of English, some differences in the use of Arabic codes (Arabic and Arabizi) were found. Unlike the assumption that conservatives would use Arabic more, the study's results showed that liberals used Arabic more frequently than conservatives who, in contrast, used Arabizi more.

Religiosity proved to make no difference in relation to the code use on Twitter. The percentage of use of different codes was almost exactly the same for all participants. These results cannot be generalized to all fields of life, but they at least can give us an indication to how polyglossic Egyptian users of Twitter act on this medium. These results can also be explained against literature (see Bassiouney, 2009) which explained that in Egypt, unlike other Arab countries like Iraq and Bahrain, both Muslims and Christians have little or no linguistic difference. Although my study is not comparing between Muslims and Christians' code choice, but the fact that religiosity resulted in no such difference is not surprising according to previous research.

Accommodation and stance. From the percentages and trends of accommodation and the choice of codes which the respondents replied with, it can be concluded that participants accommodated more when the tweet was written in one code and not in code-switching between two codes. Specifically with tweets written in English, respondents showed a great tendency to reply in English with at least 75% of them doing so. Accommodation occurred less when the tweet was written in one of the Arabic codes (Arabic or Arabizi). Since tweets written in CS by

[4] Suleiman, Y. (2003). *The Arabic language and national identity: A study in ideology*. Koinonia, Manchester: Edinburgh University Press.

default are written in two codes, it is not surprising that respondents replied with either or both of them but did not make an effort to accommodate to CS in particular.

Describing the general tendency of accommodation, one can note that participants in this study generally accommodated the original tweet writers. Reasons for accommodation, as explained by Giles and Powesland (1997), cannot be fully explained without getting back to the participants. However, since there were many more cases of disagreement than agreement expressed in the replies to the tweets, I can conclude that respondents accommodated not to reduce dissimilarities with the tweet authors but likely to ensure understanding of the message. Given that they are able to, Twitter users might use the same code used by the person they are replying to in order to make sure that what they are saying is understood by that person. Participants preferred a clearer communication rather than risking the loss of the message if another code were used. This indicates a stance of assertiveness by the participants who wish to be understood and to communicate their agreements as well as their disagreements in clear words and who openly express the (dis)similarities of their ideas to those of the tweet authors they are replying to.

Replying methods and stance. Du Bois (2007) explained that to understand stance, one has to identify who the stancetaker is, what the counter-stance is—the stance which the stancetaker is responding to—and what the object of stance is. In this study, the stancetakers are polyglossic Egyptians who are taking a stance on Twitter through responding to a tweet they have supposedly read on Twitter, and the object of stance is either the topic delivered by each of the tweets to which they are responding or the writer of the original tweet. The methods used to take these stances are outlined in the following paragraphs.

Through a qualitative analysis of six of the tweets in the DCT (see Appendix A), it was found that participants tended to frame their replies using a subjective dimension (by framing the reply around themselves or the authors of the tweets) or in an objective manner (by framing the reply around the content of the tweet). Supported by different stance markers, respondents were able to reply to the tweets and to take a stance by positioning themselves in relation to the tweet and its writer accordingly.

One conclusion that can be drawn from the data analyzed is that polyglossic Egyptians are very assertive in their replies. The analysis of the second, third, sixth, eighth, ninth, and tenth tweets showed that participants positioned themselves and expressed their alignment or lack of it quite bluntly. Participants mostly took an epistemic stance to express their opinions. They were very certain of what they were saying, and preferred to stand out and to express their opinions freely in their tweets. They preferred to take direct stances instead of hidden ones, which meant that they felt attached to the topics being discussed and wanted to express opinions regarding them. The epistemic stance was a means of claiming position of superiority and high understanding and knowledge of the truth. This was achieved through the use of boosters (very few hedges were used), sarcasm, simple present tense (to express an opinion as if stating a fact), and modals (to offer advice).

It is important to mention here the character limit on Twitter of only typing replies in not more than 140 characters contributes to such result. Such a short response likely precludes a complex, nuanced argument in favor of a quick, clear message of stance. The character limit on Twitter can explain why the stance most taken by respondents was the epistemic stance.

In many cases, this epistemic stance was even reinforced with direct evaluation whereby stancetakers evaluated the object of stance which was either the topic to which they responded or

the person to whom they are replying. By making a direct evaluative act, stancetakers were able to place a judgment on the object of stance which was either the content of the topic or the tweet author. Markers of such an act were the adjectives which they used to describe the object of stance.

Stance and identity. In this study, the medium through which the stancetakers expressed their identities is Twitter. Thus, it is very important to note that this is going to be an explanation of their identity online. Paltridge (2012) explained that there is a difference between an identity online and an identity offline; hence, we should not try to make the mistake of generalizing such findings to going beyond Twitter. Further, Paltridge (2012) has explained that one's identity depends on the context, the purpose of the discourse, the way we use language, and the way we interact with people. Bearing these points in mind, I attempt to explain the relationship between stance taken by the respondents and their identities.

It can be concluded from the results of the study that polyglossic Egyptians have an assertive identity on Twitter. Evidence of the assertiveness that they choose to display are: a) the epistemic stance which they preferred to take, which was occasionally accompanied by direct evaluation of the object of stance; b) the direct alignment or disalignment with the tweet authors; c) the heavy use of sarcasm as a marker of disagreement; d) the scarcity of hedges in their replies; and e) choosing to accommodate their code to that of the original tweet to which they are replying and hence ensuring a better communication.

Further, their online identity is apparently not related to a direct expression of feelings. As shown from the examples in chapter 4, an affective stance was never taken by any of the stancetakers in the study. On the contrary, replying was a method by which stancetakers

expressed either personal opinion in fact-like manner or evaluated the opinion of the tweet author to whom they were responding.

Implications and Conclusions

Previous research has always focused on understanding Twitter's conventions (like the hashtag and the retweets), or were more directed towards developing new software for Twitter Sentiment Analysis (TSA)[5]. Analyzing sentiments used on Twitter is a way which could help companies to understand their customers' feelings about products, politicians to understand their constituents' desires, and social organizations to understand people's opinion on debates (Pak & Paroubek, 2010). There were no studies, to my knowledge, which examined the identity of Twitter users through exploring the relationship between stance and Twitter replies. The implications of the study are:

- Polyglossic Egyptians use Twitter as a medium for free expression where they do not wish to hide their opinions. Rather, they openly take an epistemic stance marking themselves as knowledgeable and stating their opinions as facts, while expressing their alignment or disalignment. The epistemic stance they took was occasionally supported by direct evaluative acts of casting an evaluation on the tweet author to whom they were replying or on the author's opinion.
- Considering oneself politically or religiously as conservative, liberal, or neither made no difference to the choice of English code. Polyglossic Egyptians most commonly used English to reply to tweets on Twitter regardless of whether the tweet they were replying to was written in English or not. There is no clear correlation between code-choice and the political and religious ideologies of polyglossic Egyptian stancetakers on Twitter.

[5] To know more about Twitter Sentiment Analysis (TSA), read Go, Bhayani, and Huang (2009), Pak and Paroubek (2010), and Liu, Li, and Guo (2012). For full reference to these studies, check the reference list.

This appears to contradict previous claims that conservatives are more likely to use Arabic and less likely to use English as compared to liberals.

- Although the general preference for code-choice was English, polyglossic Egyptians tended to accommodate their choice of code to the code of the tweet to which they were replying. Accommodation in such cases may be related to the direct stance they took and the assertive identity they displayed, where they left no space for ambiguity and did not risk the loss of meaning.
- The stance taken by polyglossic Egyptians is the mediating factor which can help us understand the online identity of polyglossic Egyptians. The identity they display online is an assertive one showing themselves as knowledgeable people guided by facts about topics and not feelings.

Limitations of the Study

There are some limitations to this study. The first is that it all depends on self-reported data. In other words, there is no way to check the validity of the replies to the demographic section of the DCT (see Appendix A) at least. However, verifying these data is very hard to do. It requires either very close knowledge of each of the participants over a long period of time, which is not possible in this study for time and human constraints, or asking their teachers, which is also not an available option, especially with the big number of participants in this study.

Another limitation is that the replies obtained through this study are not authentic despite all the careful measures taken to attain as genuine replies as possible. Conducting a similar study on authentic replies would render even more reliable results, but this was not possible to do in this study, since it will require better knowledge of the participants.

Finally, requiring information about participants' political identity is considered sensitive information, especially at that time in Egypt with the political unrest occurring since January 25, 2011. It could be one of the reasons why almost 50% chose to identify themselves as neither liberal nor conservative.

Recommendations for Further Research

In the course of this study, some areas were identified as fertile for more future research, especially with reference to the time and focus limitations of the study. One of these areas is comparing the results of this study to other social media websites. A more thorough comparative study should be able to address this better through, for example, comparing the results of this study to how people express their stance on a blog where there is more space for using the language and is not character-limited.

I have noticed two particular aspects about the use of Twitter which could render some interesting findings if explored deeply. These are the highly excessive use of sarcasm in the replies and the reference to mothers as a way of insulting the person one is replying to. Sarcasm was frequently used in all the codes, but the reference to mothers was only done in Arabic and Arabizi.

In addition, future research as indicated in the "Limitations" section above could also conduct a similar study on authentic tweets, rather than on replies through observing how Twitter users take stance in relation to a topic, since the presence of an author to whom people are replying will no longer be a variable. In such case, stance can be linked to the code used and to dialogicality rather than to accommodation.

During the stage of collecting the data for the pilot study and the final one, and in asking students whether they used Twitter or not, many replied that they did not. Others replied that

they used to but not anymore, and others said that they had an account but they did not really use it. A particular student even offered me an advice of not conducting my study on Twitter, since it "is dying, don't you think?" These answers from different students might be indicators of the future of Twitter and some social media platforms. Because this is not a reliable indicator, it would be interesting if this can be measured somehow or looked into in more depth.

There is no literature about the code which I called Arabized English (AE). It would be interesting to examine this code further to see when, how, and why it was developed and when and why people use it. From my own personal observation of social media platforms, it is mostly used in a sarcastic way and usually along with another code. In other words, it is usually used in code-switching between AE and another code. Unlike Arabizi, people do not conduct full conversations in AE. They even quite rarely type long sentences in it. These observations need to be supported (or not) by literature to gain more reliability.

Finally, I have also noted that there were differences in the use of English and Arabizi codes between males and females. Females used English for ~56% of the replies, whereas males used it for ~41.5%. Arabizi, on the other hand, was used by males (~29.2%) more than females (~16.8%). The use of other codes was almost the same though. This, however, was not investigated further in this study but represents a fertile area for future research.

REFERENCES

Abu Elhij'a, D. (2012). Facebook written Levantine vernacular languages. *Levantine Review 1*(1), 68-105.

Androutsopoulos, J. (2006). Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics, 10*(4), 419-38.

Bassiouney, R. (2009). *Arabic sociolinguistics*. Edinburgh, UK: Edinburgh University Press.

Bassiouney, R. (2010). Identity and code-choice in the speech of educated women and men in Egypt: Evidence from talk shows. In R. Bassiouney (Ed.), *Arabic and the media,* (pp. 97-123). Leiden, Boston: Brill.

Bassiouney, R. (2012). Politicizing identity: Code choice and stance-taking during the Egyptian revolution. *Discourse and Society, 23*(2), 107-126.

Bassiouney, R. (2014). *Language and identity in modern Egypt.* Edinburgh, UK: Edinburgh University Press.

Bassiouney, R. (2015). Dialect and stance-taking by non-Egyptian celebrities in Egypt. *Open Linguistics, 1*(1), 614-633.

Boyd, D., Golder, S., & Lotan, G. (2010, January). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on System Sciences* (pp. 1-10).

Bucholtz, M. & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies, 7*(4-5), 585-614.

Burger, J. D., Henderson, J., Kim, G., & Zarrella, G. (Eds.). (2011). *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Bedford, Massachusetts.

Charles, M. (2006). The construction of stance in reporting clauses: A cross-disciplinary study of theses. *Applied Linguistics*, *27*(3), 492-518.

Conrad, S. & Biber, D. (2000). Adverbial marking of stance in speech and writing. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse*, (pp. 56-73). New York: Oxford University Press.

Doyle, G. (2012). The Hashtag's not ruining anything. *Motivated Grammar*. Retrieved from https://motivatedgrammar.wordpress.com/2012/01/03/the-hashtags-not-ruining-anything/

Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, (pp. 139-182). Philadelphia, PA: John Benjamins.

Elsadda, H., (2010). Arab women bloggers: The emergence of literary counterpublics. *Middle East Journal of Culture and Communication*, *3*(3), 312–332.

Englebretson, R. (2007). Stancetaking in discourse: An introduction. In R. Englebretson (Ed.) *Stancetaking in discourse: Subjectivity, evaluation, interaction*, (pp. 1-26). Philadelphia, PA: John Benjamins.

Ferguson, C. A. (1959). Diglossia. *Language and Social Structures*, *15*(2), 232-251.

Fishman, J. A. (1967). Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues*, *23*(2), 29-38.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In Giles H, Coupland, J., & Coupland, N. (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1-68). Cambridge: Cambridge University Press.

Giles, H., & Powesland, P. (1997). Accommodation theory. In N. Coupland & A. Jaworski (Eds.) *Sociolinguistics: A reader* (pp. 232-239). Essex, UK: Macmillan Education.

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, 1, 12.

Haddington, P. (2007). Positioning and alignment as activities of stancetaking in news interviews. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 283-317). Philadelphia, PA: John Benjamins.

Halim, N. S., & Marlyna M. (2014). The functions of code-switching in Facebook Interactions. *Procedia-Social and Behavioral Sciences, 118*: 126–33.

Hansen, D., Shneiderman, B., & Smith, M. A. (2011). *Analyzing Social Media Networks with NodeXL: Insights from a connected world.* Burlington, MA: Morgan Kaufmann.

Herring, S. C. (2007). A faceted classification scheme for computer-mediated discourse. *Language@ Internet*, *4*(1), 1-37.

Herring, S. C. (2011). Discourse in Web 2.0: Familiar, reconfigured, and emergent. In D. Tannen & A. M. Trester (Eds.), *Discourse 2.0: Language and new media* (1-25). Washington, DC: Georgetown University Press.

Holmes, J. (2013). An introduction to sociolinguistics. New York, NY: Routledge.

Hong, L., Convertino, G., & Chi, E. H. (2011). Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media. Menlo Park, California: The AAAI Press.

Hunston, S. (2007). Using a corpus to investigate stance quantitatively and qualitatively. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 27-48). Philadelphia, PA: John Benjamins.

Hurreiz, S. H. (1975). Arabic in the Sudan: Current problems and future needs. *Language Planning Newsletter*, *1*(4), 1-3.

Hyland, Ken. 2005. Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173-192.

Jaffe, A. (2009). Stance: Sociolinguistic perspectives. In M. Bucholtz (Ed.), *From stance to style: Gender, interaction, and indexicality in Mexican immigrant youth slang* (pp. 1-49). Oxford, UK: Oxford University Press.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. San Jose, California.

Johnstone, B. (2007). Linking identity and dialect through stancetaking. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 49-68). Philadelphia, PA: John Benjamins.

Johnstone, B. (2008). *Discourse analysis*. Malden, MA: Blackwell.

Keisanen, T. (2007). Stancetaking as an interactional activity: Challenging the prior speaker. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 253-281). Philadelphia, PA: John Benjamins.

Kosoff, Z. (2014). Code-switching in Egyptian Arabic: A sociolinguistic analysis of Twitter. *Al-ʿArabiyya: Journal of the American Association of Teachers of Arabic*, *47*, 83-99.

Krishnamurthy, B., Gill, P., & Arlitt, M. (2008). *Proceedings of the first workshop on Online social networks*. Seattle, Washington.

Liu, K., Li, W., & Guo, M. (Eds.). (2012). Proceedings from *19th International Conference, DASFAA 2014*. Chennai, India: Springer

McCann R., Ota H., Giles H., & Caraker R. (2003). Accommodation and nonaccommodation across the lifespan: Perspectives from Thailand, Japan, and the United States of America. *Communication Reports*, *16*(2), 69-91.

Merchant, G. (2001). Teenagers in cyberspace: An investigation of language use and language change in internet chatrooms. *Journal of Research in Reading, 24*(3), 293-306.

Muhammed, R. (2013). *Language used by Egyptians on Facebook and January 25threvolution: A descriptive study September 2010-Septmber 2011*. (Unpublished master's thesis). The American University in Cairo, Cairo, Egypt.

Paltridge, B. (2012). *Discourse analysis: An introduction*. Chennai, India: Bloomsbury.

Platt, J. (1977). A model for polyglossia and multilingualism (with special reference to Singapore and Malaysia). *Language in Society*, *6*(3), 261-378.

Page, R., Barton, D., Unger, J. W., & Zappavigna, M. (2014). *Researching language and Social Media: A student guide.* New York, NY: Routledge.

Pak, A., & Paroubek, P. (2010). Proceedings from *The Seventh conference on International Language Resources and Evaluation.* Orsay Cedex, France.

Rauniomaa, M. (2007). Stance markers in spoken Finnish: Minun mielestä and minusta in assessments. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 221-252). Philadelphia, PA: John Benjamins.

Silver, M. (2003). The stance of stance: A critical look at ways stance is expressed and modeled in academic discourse. *Journal of English for Academic Purposes, 2*(4), 359-374.

Tannen, D. (2013). The medium is the metamessage: Conversational style in new media interaction. In D. Tannen & A. M. Trester (Eds.), *Discourse 2.0: Language and new media* (99-117). Washington, DC: Georgetown University Press.

West, L., & Trester, A. M. (2013). Facework on Facebook: Conversation on Social Media. In D. Tannen & A. M. Trester (Eds.), *Discourse 2.0: Language and new media* (133-154). Washington, DC: Georgetown University Press.

Yaghan, M. A. 2008. "Arabizi": A contemporary style of Arabic slang. *Design Issues*, *24*(4), 39-52.

Zappavigna, M. (2012*). Discourse of Twitter and social media: How we use language to create affiliation on the web.* Chennai, India: Bloomsbury.

Appendix A

DCT

Demographic Information

Answer the following questions:

1. What is your nationality?
    a) Egyptian
    b) Other. (Please specify) …….
2. You are
    a) Male
    b) Female
3. What is your age?
    a) 16 – 19 years old
    b) 20 – 23 years old
    c) 24 – 27 years old
    d) 28 or older
4. What do you think your native language is?
    a) Arabic
    b) English
    c) Other. (Please specify) …….
5. How would you describe your command of Arabic?
    a) Weak
    b) Intermediate
    c) Advanced
6. How would you describe your command of English?

a) Weak

b) Intermediate

c) Advanced

7. What type of school did you attend?

a) Public (Arabic)

b) Public (Experimental)

c) Private

d) International

8. You politically consider yourself

a) Liberal

b) Conservative

c) None of the above

9. You religiously consider yourself

a) Religious

b) Religious and liberal

c) Non-religious

d) None of the above

10. Do you have a Twitter account?

a) Yes

b) No

11. When did you join Twitter?

a) 2015 or 2016

b) 2013 or 2014

c) 2011 or 2012

d) Before 2011

12. How frequently do you use Twitter?

a) Rarely

b) Monthly

c) Weekly

d) Daily

13. How many tweets have you posted so far?

a) 0 – 100

b) 101 – 1,000

c) 1,001 – 10,000

d) More than 10,000

14. How many followers do you have so far?

a) 0 – 100

b) 100 – 500

c) 501 – 1,000

d) 1,000 or more

15. What is your Twitter handle (e.g. @abc123)? **(OPTIONAL)**

Tweeting

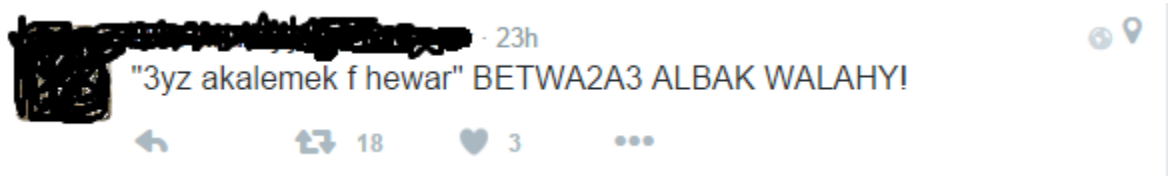
- The following are screenshots of ten real tweets taken from different random accounts on Twitter. If you were to write a tweet in reply to each, what would it be? Write your reply underneath each of them.

- Usernames and profile pictures are hidden to preserve users' anonymity.

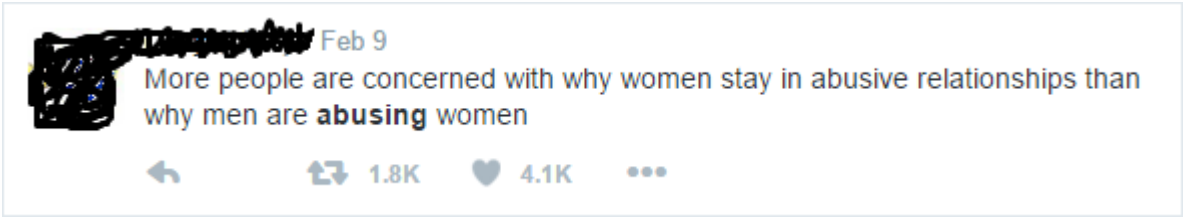
- You are encouraged to reply in any language you normally tweet in, like Arabic, English, a mix of both, or Franco-Arab.

Please bear in mind that your replies should NOT exceed 140 characters.

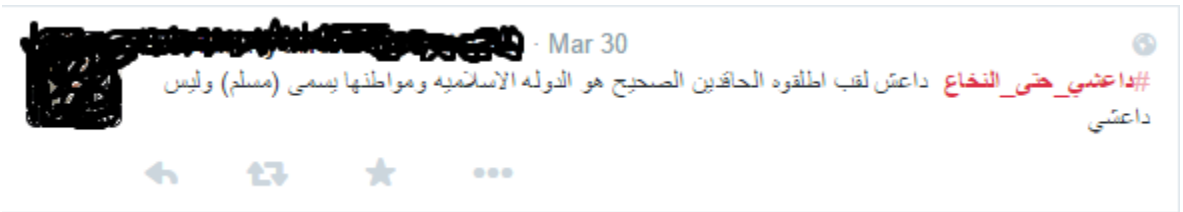
1.



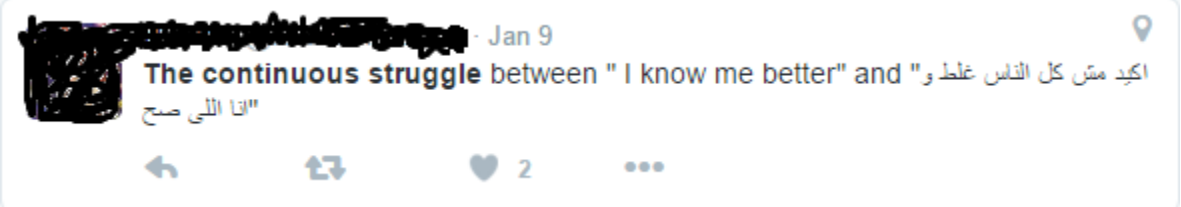
2.



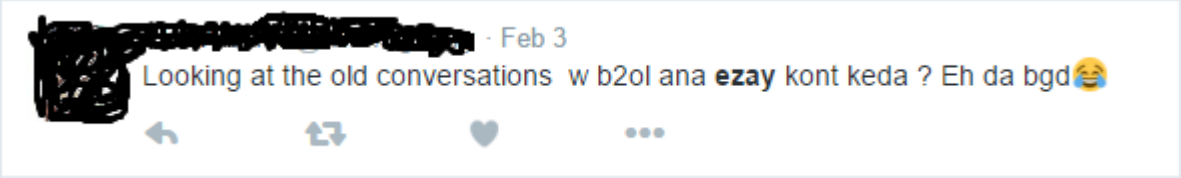
3.



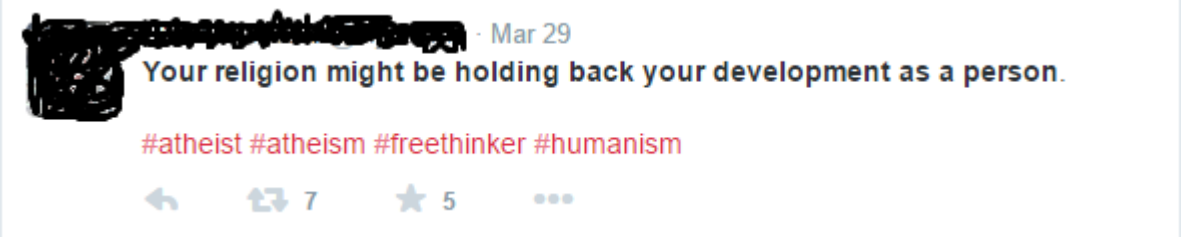
4.


· Jan 9
The continuous struggle between " I know me better" and "اكيد مش كل الناس غلط و انا اللى صح"
2

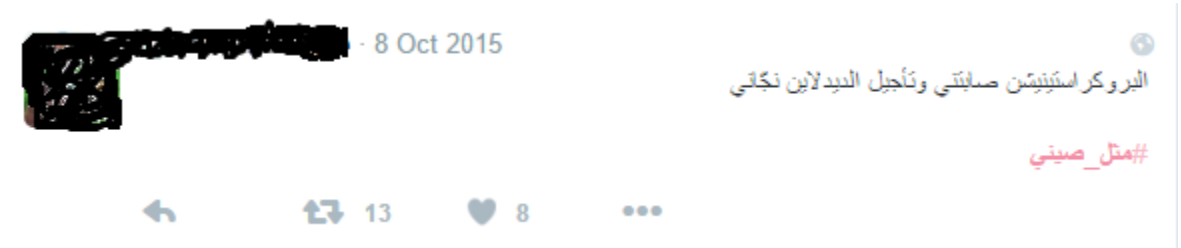
5.


· Feb 3
Looking at the old conversations w b2ol ana ezay kont keda ? Eh da bgd😂

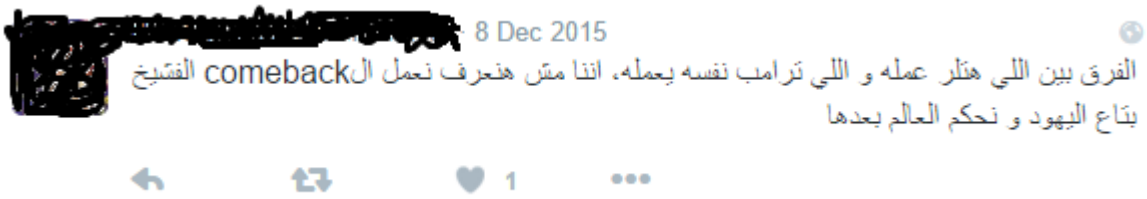
6.


· Mar 29
Your religion might be holding back your development as a person.
#atheist #atheism #freethinker #humanism
7
5

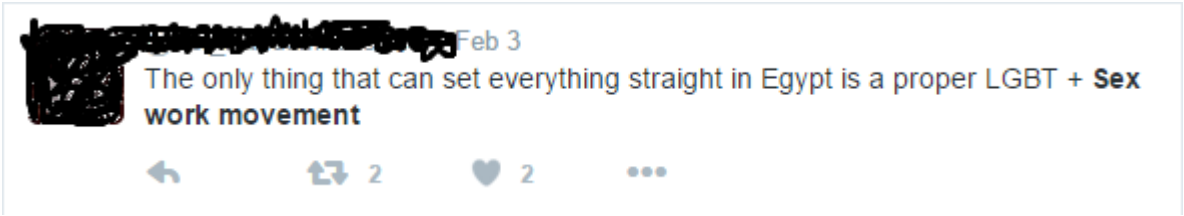
7.


· 8 Oct 2015
البروكراستينيشن صاحبتي وتأجيل الديدلاين نجّاني
#متل_صيني
13
8

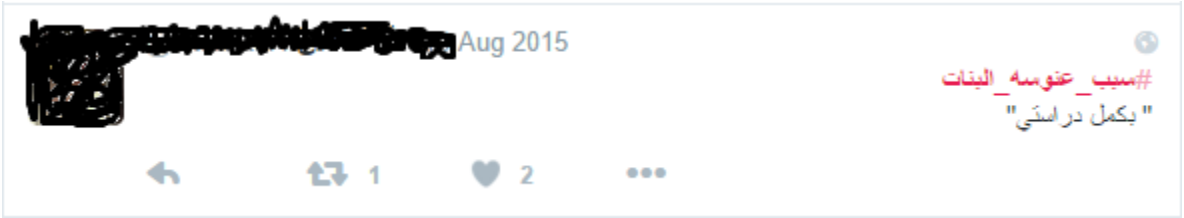
8.


8 Dec 2015
الفرق بين اللي هنكر عمله و اللي ترامب نفسه يعمله، اننا مش هنعرف نعمل الcomeback الفشيخ
بتاع اليهود و نحكم العالم بعدها
1

9.


Feb 3
The only thing that can set everything straight in Egypt is a proper LGBT + **Sex work movement**
2
2

10.


Aug 2015
#سيب_عنوسه_البنات
" بكمل دراستي"
1
2

Appendix B

Translation of DCT Tweets

Tweet #	Tweet	Translation
1	“3yz akalemek f hewar” BETWA2A3 ALBAK WALAHY!	“We need to talk” can give you a heart attack!
3	#داعشي_حتي_النخاع داعش لقب اطلقوه الحاقدين الصحيح يسمي الدولة الاسلامية و مواطنها يسمي (مسلم) و ليس داعشي	#Daeshi_to_the_core Daesh was coined by haters. The correct term is the Islamic State, and its citizen is called a (Muslim) not a Daeshi
4	The continuous struggle between “I know me better” and "اكيد مش كل الناس غلط و انا اللي صح"	The continuous struggle between “I know me better” and “they can’t be all wrong and I’m right”
5	Looking at the old conversations w b2ol ana ezay kont keda? Eh da bgd 😂	Looking at the old conversations, I was like: What was I thinking? What was that? 😂
7	البروكراستينيشن صابتني و تأجيل الديدلاين نجّاني #مثل_صيني	Procrastination was gonna get me but extending the deadline saved me #Chinese_proverb
8	الفرق بين اللي هتلر عمله و اللي ترامب نفسه يعمله، اننا مش هنعرف نعمل الcomeback الفشيخ بتاع اليهود و نحكم العالم بعدها	The difference between what Hitler did and what Trump wants to do is that we won’t be able to make the same dramatic comeback the Jewish did and rule the world afterwards
10	#سبب_عنوسه_البنات "بكمل دراستي"	#cause_of_female_spinsterhood “Finishing my education”

Appendix C

IRB Approval of Study

CASE #2015-2016-092

THE AMERICAN UNIVERSITY IN CAIRO
INSTITUTIONAL REVIEW BOARD

To: Sahar Mashhour

Cc: Sara Tarek

From: Atta Gebril, Chair of the IRB

Date: Feb 11, 2016

Re: Approval of study

This is to inform you that I reviewed your revised research proposal entitled "**Code Choice on Twitter: How Stance-taking Reflects the Identity of Polyglossic Egyptian Users**" and determined that it required consultation with the IRB under the "expedited" heading. As you are aware, the members of the IRB suggested certain revisions to the original proposal, but your new version addresses these concerns successfully. The revised proposal used appropriate procedures to minimize risks to human subjects and that adequate provision was made for confidentiality and data anonymity of participants in any published record. I believe you will also make adequate provision for obtaining informed consent of the participants.

This approval letter was issued under the assumption that you have not started data collection for your research project. Any data collected before receiving this letter could not be used since this is a violation of the IRB policy.

Please note that IRB approval does not automatically ensure approval by CAPMAS, an Egyptian government agency responsible for approving some types of off-campus research. CAPMAS issues are handled at AUC by the office of the University Counsellor, Dr. Amr Salama. The IRB is not in a position to offer any opinion on CAPMAS issues, and takes no responsibility for obtaining CAPMAS approval.

This approval is valid for only one year. In case you have not finished data collection within a year, you need to apply for an extension.

Thank you and good luck.

Atta Gebril

Dr. Atta Gebril
IRB chair, The American University in Cairo
2046 HUSS Building
T: 02-26151919
Email: agebril@aucegypt.edu

Institutional Review Board
The American University in Cairo
AUC Avenue, P.O. Box 74
New Cairo 11835, Egypt.
tel 20.2.2615.1000
fax 20.2.27957565
Email: aucirb@aucegypt.edu

Appendix D

Consent Form for Participation

THE AMERICAN UNIVERSITY IN CAIRO
INSTITUTIONAL REVIEW BOARD

Documentation of Informed Consent for Participation in Research Study

Project Title: Code Choice on Twitter: How Stance-taking Reflects the Identity of Polyglossic Egyptian Users

Principal Investigator: Name: Sahar Mashhour
E-mail: saharm@aucegypt.edu

You are being asked to participate in a research study which will constitute part of my MA thesis. The purpose of the research is to investigate how Egyptians use Twitter as a medium for self-expression, and the findings may be published, presented, or both. The expected duration of your participation is 10 to 15 minutes.

The procedures of the research will be as follows: data from participants will be collected then analyzed by the researcher.

There will not be any risks or discomforts associated with this research.

There will not be benefits to you from this research.

The information you provide for purposes of this research is confidential.

Questions about the research or your rights should be directed to Sahar Mashhour at saharm@aucegypt.edu.

Participation in this study is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time without penalty or the loss of benefits to which you are otherwise entitled.

By clicking "Next," you agree to participate in this study.

Thank you very much for your willingness to help!