The Dialectal Shared Linguistic Features of the Diverse Arabic Dialects Resulting a Lingua Franca Dialect of Arabic

Ameer Ali Hussein Al Majtoomy

Lecturer M.A. linguistics in University of Kufa, Najaf, Iraq

Abstract: *Arabic is a very morphologically rich Semitic language so it has several dialects across the Arab homeland, the paper focuses on the similarities and differences between the Arabic dialects. The research tries to find out the most shares features among the Arabic dialects and decide a dialect a lingua franca of Arabic language to be used and taught as for learners of Arabic. This lingua franca dialect could be used outside the Arabic countries.*

Keywords: MSA, dialect, lingua franca, Levantine dialect, Gulf dialect, Maghreb dialect

1. Introduction

This topic addressed a very important phenomenon in western society where immigrants of different Arabic dialects can communicate easily in spite of the difference between those dialects. The main problem of this project is to shed the light on the similarities between the Arabic dialects which make it possible for those speakers understand each other. The second question is how these diverse dialects make a speech Arabic community in the western society, is it because of the syntactic, semantic and phonological similaritiesbetween the Arabic dialects or/ and also because of the shared sociolinguistic communicative competence of those speakers who belong to the same mother language(Modern Standard Arabic) which help the speakers to re-contextualize other dialects.

By using some qualitative methods such as surveys and questionnaires, the participants of this project who are the immigrants outside the Arab countries in addition to other Arabic speakers such as students, tourists and the like, will help us conclude the shared features of Arabic dialects and answer all the other research questions.

2. Research Questions

1) What are the shared features of the diverse Arabicdialects according to the phonological, morphological, syntactic and semantic levels?
2) How these similarities help the speakers of diverse Arabic dialects to understand each other?
3) Can these dialects compose a speech community of Arabic in the western society?
4) Is there any remark for Arabic to be a diglossic language abroad?
5) Can we determine for a specific dialect to be the Lingua Franca of all Arabic dialects? If possible, what makes that specific dialect to be like that?
6) What is the role of the sociolinguistic communicative competence to re-contextualize other dialects?

3. Aims of the study

This study aims:
1) To identify the shared linguistic features according to the phonological,morphological, syntactic, and semantic levels.
2) To know how the similarities or shared features help the speakers of Diverse Arabic dialects to understand each other.
3) To decide whether these diverse dialects of Arabic compose a speech community.
4) To specify which Arabic dialect could be considered the Lingua Franca of Arabic.
5) To capture the importance of the sociolinguistic communicative competence of the diverse Arabic speakers to re-contextualize other dialects.

4. Research Hypotheses

The research hypothesizes the following:
1) All the Arabic dialects share many featuresas they belong to one mother language " Arabic"
2) Those similarities help the speakers of diverse Arabic dialects to understand each other.
3) The dialects of Arabic in global cities such as London or Manchester in the UK make a speech community of Arabic.
4) Arabic is a diglossic language whether it is used in Arabic countries or other global cities.
5) There is a specific dialect which could be regarded as a Lingua Franca of Arabic.
6) The sociolinguistic communicative competence of the speakers has a great importance to re-contextualize other dialects.

Value of the Study
The value of this study fall into two ways, Firstly this study will help the Arabic immigrant speakers of Diverse Arabic Dialects to adopt one dominant dialect as a lingua Franca dialect which will be intelligible for all of them, Secondly which is the most important is that this study will give the chance to schools and institutions to teach the most

intelligible understandable Arabic dialect to AFL (Arabic as a Foreign Language) Learners.

5. Literature Review

Many studies and works have tackled Arabic dialects and shown the classification of Arabic dialects into several types according to regional distribution, those studies also have shown the similarities and differencesbetween these types, the results of some of these works will be stated, some contextual examples will be given in addition to some experiments and the approaches used to analyze the data.

The formal form of Arabic language is Modern standard Arabic (MSA) which is originally the language of the Holy Qura'an. MSA is formally used now in communications and declarations and office affairs by the governments of Arabic countries, and it is also the language of schools, universities, institutions and any other academic foundations. On the other hand, Arabic dialects are used by Arabic speakers in everyday life, because of the highly usage of these dialects, a lot of art or literary works are produce in these dialects such as movies, series, stories, folktales and public blank poetry(Biadsy et al 2009, p: 54).

Arabic language was firstly formulated in Arabian Peninsula but by the spread of Islam, the Arabic language is spread to the neighboring countries what so called today the Arab world. The language of Arabs in the Arabian Peninsula was only the formal Arabic, but after the spread of Islam out of that peninsula, Arabic has undergone many transformations to have different dialects started exactly from the 3rd decade of the seventh century. The main division of the Arabic dialects is Eastern and Western dialects , Eastern dialects include the dialects of Saudi Arabia, Kuwait, Qatar, Bahrain, United Arab Emirates, Yemen , Oman along with Iraq, Syria, Lebanon, Palestine, Jordan, Egypt and Sudan. The western dialects are found in Libya, Tunisia, Algeria, Morocco and Mauritania. The subdivision of Arabic dialects according to the region results five types:

1) Iraqi dialects
This dialect s spoken only in Iraq, one of the specific features of this dialect is changing the phoneme / ɋ/ into voiced / g / or unvoiced / k /.

2) The Gulf dialect
This dialect is spoken in all the gulf countries (Persian) Gulf such as Kuwait, Qatar, Saudi Arabia, UAE and Oman. The specific feature of this dialect is that the phoneme / dʒ/ changed to / j /, and the phoneme / ɋ / changed to voiced / g /.

3) Levantine dialect
The dialect is spoken by the people in the Mediterranean east countries such as Jordan, Syria, Lebanon, and Palestine. This dialect is sub-divided into six sub-dialects.
a) Lebanese dialect.
b) Central Syrian.
c) Northern dialect.
d) Rural Palestinian.
e) Urban Palestinian.
f) Bedouin Palestinian.

4) Egyptian dialect
This dialect is spoken by the Egyptian people, this dialect is the most wide spread dialect because of the Egyptian films, series and programmes which are so popular in the Arabic media. This is why so many American and European schools taught Egyptian dialect. In this dialect / θ/ changes to / d/ and /dʒ/ to / g/ such as / dʒəmi:l (beautiful) changed to / gəmi:l/.

5) Maghreb dialect
Maghreb which means "west" refers to the countries Morocco, Tunisia, Algeria and western Libya. The specific feature of this dialect is the influence of the Spanish and French languages on this dialect so it includes a lot of words from these two languages as parts of the Maghreb Dialects (Alorifi 2008, P: 7-9)
Here is a table of examples of some Arabic dialects:

Table 1: A table comparing the translations for one sentence in the Multidialectal Arabic Corus. Egyptian Arabic is the original sentence which was translated to MSA, English and the other dialects.

Dialect/Language	Example
English	*Because you are a personality that I can not describe.*
Modern Standard Arabic	لأنك شخصية لا أستطيع وصفها. *lÂnk šxSyħ lA ÂstTyς wSfhA.*
Egyptian Arabic	لأنك شخصية وبجد مش همرف أوصفها. *lÂnk šxSyħ wbjd mš hςrf ÂwSfhA.*
Syrian Arabic	لأنك شخصية وعنجد ما رح أعرف أوصفها. *lÂnk šxSyħ wςnjd mA rH Âςrf ÂwSfhA.*
Jordanian Arabic	انت جد شخصية مستحيل اقدر اوصفه *Ant jd šxSyħ mstHyl Aqdr AwSfhA.*
Palestinian Arabic	عن جد ماشاء الله عليك شخصيتك ما بتنوصف. *ςn jd mA šA' Allh ςlyk šxSytk mA btnwSf.*
Tunisian Arabic	على خاطرك شخصية بلحق منجمش نوصفها. *ςlý xATrk šxSyħ blHq mnjmš nwSfhA.*

(Bouamor et al 2015, p: 1242)

Other examplesof Levantine Dialects, Egyptian and Maghreb can be stated here as:
a) Lam jaʃtarikamâlʈawilatandʒadi:datan (MSA) لم يشتري كمال طاولة جديدة
b) Didn't buy Kamal table new (gloss)
c) kamâlmaʃtarâʃʈarabezagidi:da (Egyptian)
d) kamâlmaʃtarâʃʈawileʒidi:de (Levantine)
e) kamâlmaʃrâʃmidaʒidi:da (Maghreb)
f) Kamal not bought table new (dialectal gloss)
(Diab&Habash 2014)

By the shape of tables some other examples of MSA, Levantine and Egyptian will be given as follows:

Tables 2

Source (*EGY*):	انت بتعمل له اعلان ولا ايه؟!!
Transliteration:	*Ant btEml lh AElAn wlA Ayh?!!*
MSA-Sys. Output:	You are working for a declaration and not?
Dial-Sys. Output:	You are making the advertisement for him or what?
Reference:	Are you promoting it or what?!!
Source (*EGY*):	نفسي اطمئن عليه بعد ما شاف الصوره دي
Transliteration:	*nfsY Atm}n Elyh bEd mA SAf AlSwrh dy*
MSA-Sys. Output:	Myself feel to see this image.
Dial-Sys. Output:	I wish to check on him after he saw this picture.
Reference:	I wish to be sure that he is fine after he saw this images
Source (*LEV*):	لهيك الجو كتيير كوول
Transliteration:	*lhyk Aljw ktyyyr kwwwl*
MSA-Sys. Output:	God you the atmosphere.
Dial-Sys. Output:	this is why the weather is so cool
Reference:	This is why the weather is so cool
Source (*LEV*):	طول بالك عم نمزح
Transliteration:	*Twl bAlk Em nmzH*
MSA-Sys. Output:	Do you think about a joke long.
Dial-Sys. Output:	Calm down we are kidding
Reference:	calm down, we are kidding

Tables 3

Source (*EGY*):	قالتلة طب تعالي نعد .
Transliteration:	*qAltlp Tb tEAlY nEd ,*
MSA-Sys. Output:	**Medicine** almighty promise.
Dial-Sys. Output:	She said, **OK**, come and then
Reference:	She told him, **OK**, lets count them .
Source (*LEV*):	فبقرا وأحيانا بقضيها مع رفقاتي
Transliteration:	*fbqrA w>HyAnA bqDyhA Em >tslY mE rfqAty*
MSA-Sys. Output:	I read and sometimes with go with my **uncle**.
Dial-Sys. Output:	So I read, and sometimes I spend **trying** to make my self comfort with my friends
Reference:	So i study and sometimes I spend the time having **fun** with my friends
Source (*LEV*):	الله يسامحكن هلق كل واحد طالب قرب بيكون بدو عروس
Transliteration:	*Allh ysAmHkn hlq kl wAHd TAlb qrb bykwn bdw Erws*
MSA-Sys. Output:	God now each student near the **Bedouin** bride.
Dial-Sys. Output:	God forgive you, each one is a close student would **want** the bride
Reference:	God forgive you. Is every one asking to be close, **want** a bride!

An interesting table of Arabic dialects examples can be stated from Wikipedia:

Table 4

Variety	I love reading a lot.	When I went to the library,	I only found this old book.	I wanted to read a book about the history of women in France.
Modern Standard Arabic	ʾanā ʾuḥibbu l-qirāʾata kaṯīran	ʾindamā ḏahabtu ʾila l-maktabati	lam ʾaǧid siwā hāḏā l-kitābi l-qadīm	kuntu ʾurīdu an ʾaqraʾa kitāban ʾan tārīḫi l-marʾah fī-farānsā
Tunisian	āna nḥebb nǎqra barʃa	wǎqtelli raḥt l-əl-mǎktba	ma-lqīt kān ha-l-ktāb lə-qdīm	kŭnt nbyīt nǎqra ktāb ʕla tārīx lə-mra fi frānsa
Algerian	āna nḥebb nəqra bəzzāf	wǎqtelli raḥt l-əl-mǎktba	ma-lqīt γīr hād lə-ktāb lə-qdīm	kŭnt nbyīt nəqra ktāb ʕla tārīx lə-mra fi frānsa
Moroccan	ana ka nbri bezzaf nəqra	foqach mchit lmketaba	ma-lqīt-ʃ mən-γīr hād lə-ktāb lə-qdīm	kŭnt bāγi nəqra ktāb ʕla tārīx lə-mra fə-fṛansa
Egyptian	ana baḥebb el-ʔerāya ʔawi	lamma roḥt el-maktaba	ma-lʔet-ʃ ella l-ketāb el-ʔadīm da	ana kont ʕāyez aʔra ketāb ʕan tarīx es-settāt fe faransa
Lebanese	ktīr bḥibb il-ʔirēye	lamma reḥit ʕal-maktebe	ma lʔēt illa hal-i-ktēb li-ʔdīm	kēn beddi ʔra ktēb ʕan tērīx l-mara b-frēnse
Mesopotamian (Baghdadi)	āni aḥibb el-qrāya kulliʃ	min reḥit il-maktaba	ma liget ḥafa hāδa l-iktāb el-ʔāskī	redet aqra ktāb ʕan tarīx l-imrayyāt eb-fransa
Kuwaiti	ʔāna wāyed aḥibb aqrā kulliʃ	lamman reḥt el-maktaba	ma liget illa hal ketāb al-qadīm	kent abī aqra ketāb ʕan tarīx el-ḥarīm eb-fransa
Hejazi	ana marra aḥubb al-girāya	lamma ruḥt al-maktaba	ma ligīt γēr hāda al-kitāb al-gadīm	kunt abγa agra kitāb ʕan tārīx al-ḥarīm fi faransa

Another table by (Zaidan&Callison 2014) can illustrate the similarities and differences between the MSA and the three Arabic dialects: Levantine, Gulf and Egyptian as follows:

Table 5

English	MSA	LEV	GLF	EGY
Book	*ktAb*	*ktAb*	*ktAb*	*ktAb*
Year	*snħ*	*snħ*	*snħ*	*snħ*
Money	*nqwd*	*mSAry*	*flws*	*flws*
Come on!	*hyA!*	*ylA!*	*ylA!*	*ylA!*
I want	*Aryd*	*bdy*	*Abγý*	*ςAyz*
Now	*AlĀn*	*hlq*	*AlHyn*	*dlwqt*
When?	*mtý?*	*Aymtý?*	*mtý?*	*Amtý?*
What?	*mAðA?*	*Ayš?*	*wš?*	*Ayh?*
I drink	*Âšrb*	*bšrb*	*Ašrb*	*bšrb*
He drinks	*yšrb*	*bšrb*	*yšrb*	*byšrb*
We drink	*nšrb*	*bnšrb*	*nšrb*	*bnšrb*

By collecting around 2000 sentences covering MSA, Egyptian, Tunisian, Jordanian, Palestinian and Syrian dialects, Bouamor (2015) stated that there are some similarities between these dialects and she proved that the Egyptian is the most familiar understandable dialect to all the speakers of other dialects, but there is a little bias towards the chosen dialect which is (Egyptian), it is chosen as a starting point.

Alorifi(2008) concluded that a lot of similarities exist between the Arabic dialects; this study uses certain models to identify the specific parts or words of the Arabic dialect such as Gaussian Mixture models and Hidden Markov models. These models are used to analyze the data, for example Hidden Markov models identify the common Arabic sound and the unique sounds of certain dialects.

Biadsy(2009) made some experiments to analyze and identify Arabic dialects; these experiments are applied to evaluate the parallel PRLM approach on distinguishing Arabic dialects. Three experiments have been conducted; the first one is to compare the system of this study with the one of Alorifi(2008) on the same dialects Gulf and Egyptian. The second experiment is conducted to identify or classify the four Arabic dialects. The third one is conducted to include the MSA in five level classifications.

Sadat (2014) has also conducted some experiments to analyze and identify Arabic dialects. Three experiments are conducted by using uni-gram, di-gram, and tri-gram character for each experience based on either the Markov language model or Naive Bayes classifiers. The data is collected from the Arabic countries as shown in figure one. This figure also shows the distribution and variation in Asia and Africa.
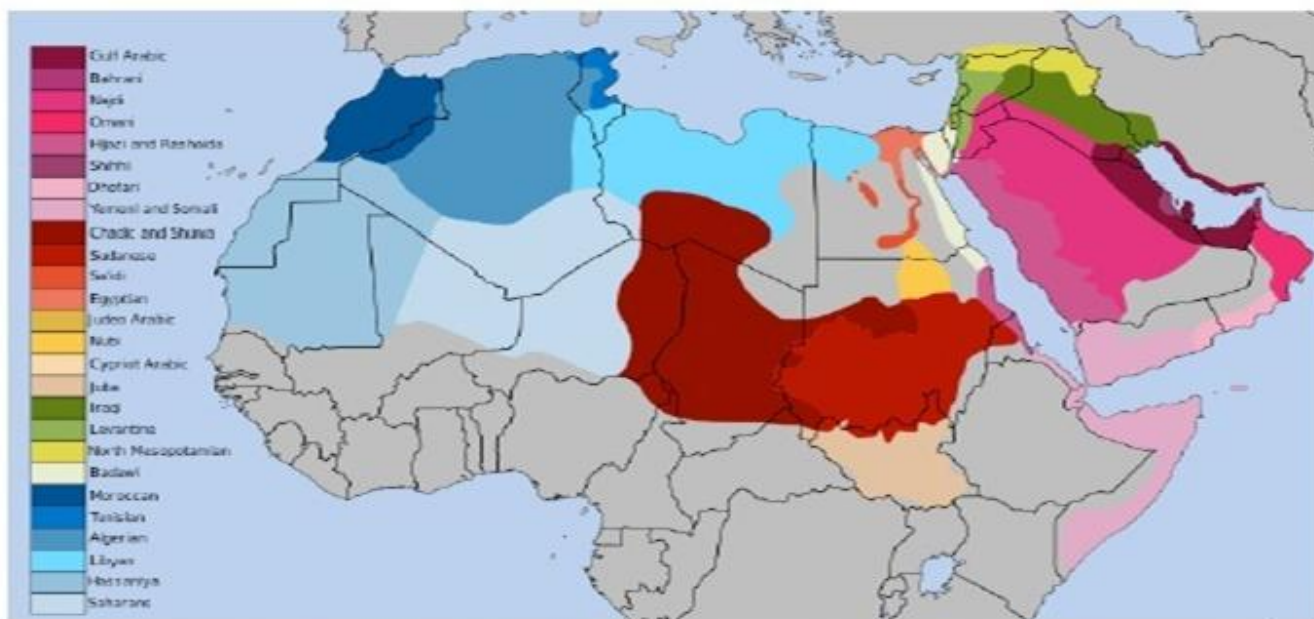

Figure 1: Arabic dialects distribution and variation across Asia

Figure 2 shows the different group divisions with their countries. Figure 3 presents a comparative picture of the accuracy dialect by using the three different character n-gram Markov language model. In these experiments, accuracy has been considered as a proportion of true identification test data, and the F- measure as a balanced mean between precision and recall. Figure 3 also shows that the character-based uni-gram distribution helps to identify two dialects, the Mauritanian and the Moroccan with an overall F-measure of 60% and an overall accuracy of 96%. The bi-gram distribution of two characters affix helps recognize four dialects, Mauritanians, Moroccan, Tunisian and Qatari with an overall F-measure of 70% and an overall accuracy of 97%. Finally tri-gram distribution of three characters affix helps recognize four dialects: the Mauritanian, the Tunisian, the Qatari and the Kuwaiti, with an overallF-measure of 73% and an overall accuracy of 98%. The comparative results of these experiments show that the based character tri- gram and bi-gram distribution

have performed better than the uni-gram distribution for most dialects.

Zbib et al (2012) also refers to the six groups of Arabic dialect according to six regions as illustrated in Figure 2.
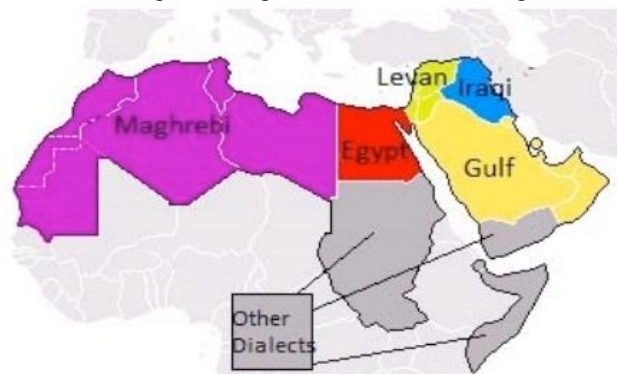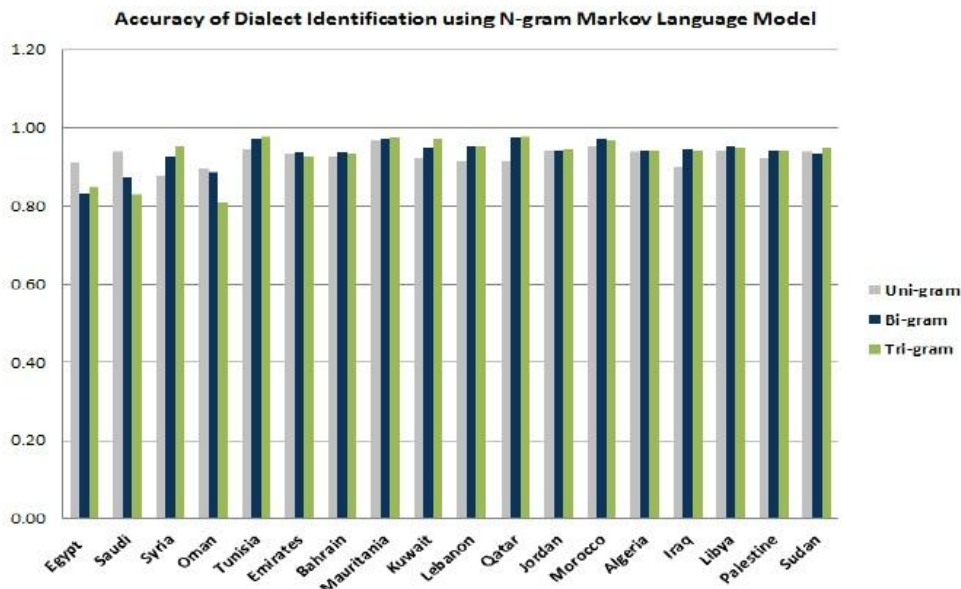

Figure 2: Division of Arabic dialects in six groups/ division

Figure 3: Accuracies on character- based N-gramMarkov language models for 18 countries

Bakir (2008) concluded that there are a lot of shared semantic items between the Iraqi dialect and the UAE dialect(a type of Gulf dialect), he pointed that those lexical items of the two dialects behave in the same way according to their relation to standard Arabic. His study also states that the dialects of east and north eastern coast Gulf countries share a lot of items.

Finally let's not forget to say that Arabic is a diglossic language which means MSA and its dialects are used side by side in the same country or region each for its specific purpose. A lot of works and studies have proved this phenomenon for example Jabbari (2012) has proved that MSA and Egyptian colloquial Arabic differa lot in the levels of phonology, morphology, lexicon and syntax. The study also proved that both of them are used in Egypt in a diglossic way but they are totally unintelligible.

6. Conclusions

At the end of this research, th following points could be concluded.
1) All the Arabic dialect share common features that make it easy for the Arabic speakers of different Arab countries understand each other.
2) Due to the similar sociolinguistic competence and the mother language (Arabic), speakers of different dialects have relative mutual intelligibility to speak and understand Arabic dialects other than their own dialect.
3) The Iraqi dialect and UAE dialect share very much in common.
4) Among the Arabic dialects, a shared dialect can be selected to be the lingua Franca of Arabic that can be used outside Arabic countries and it could be taught as a second language for non-Arabic learners.

References

[1] Alorifi, F. S. 2008, *"Automatic Identification of Arabic Dialects using Hidden Markov Models",* University of Pittsburg, Pennsylvania, USA, PH.D dissertation.

[2] Bakir, k. H. 2008, *"Lexical Similarities and Differences between Iraqi and UAE Arabic dialects"*, Journal of the College of Arts, University of Basra, Iraq. Volume 47, PP1-10.

[3] Biadsy, F , Hirschberg, J , Habash, N 2009, *" Spoken Arabic Dialects identification using phonotactic modeling"*in Semitic 09, proceedings of the EACL workshop on computational Approaches to Semitic Language, USA, PP.53-61

[4] Bouamor, H., Habash, N. ,Oflazer, K. 2013 *" A multidialectal parallel Corpus of Arabic"*, Ninth international conference on language resources and evaluation, Reykjavik, Iceland.

[5] Diab, M ,Habash, N. 2014, *" Natural Language processing of Arabic and its dialects" ,* EMNLP Conference, Doha, Qatar.

[6] https://en.wikipedia.org/wiki/Varieties_of_Arabic

[7] Jabbari, M. J. 2012, *"Diglossia in Arabic a comparative Study of the Modern Standard Arabic and Egyptian Colloquial Arabic",* Global Journal of Human Social Science, USA. Volume 12, Issue 8.

[8] Sadat, F. ,Kazimi, F. , &Farzindar, A. 2014, *" Automatic Identification of Arabic dialect in social Media",* Proceedings of SoMeRA: International Workshop on Social Media Retrieval and Analysis,New York, USA. PP. 35-40

[9] Zaidan, O. F. ,Callison, C. 2014, *" Arabic dialect Identification"* Journal of Computational linguistics, MIT, Cambridge, MA, USA. Volume 40, Issue 1, PP. 171-202.

[10] Zbib, R , Malchiodi, E , Devlin, Stallard, D. , et al 2012, *"Machine Translation of Arabic Dialects"*in NAACL HLT '12, proceedings of The conference of the North American Chapter of the Association for computational linguistics: Human language Technologies" , Stroudsburg, PA , USA. PP.49-59.