Towards Developing Algerian Arabic Orthography
Corpus-based study of Facebook Messages of Students at the Department of English, Constantine University

<div dir="rtl">

نحو تطوير أبجدية اللهجة الجزائرية العامية

تحليل لرسائل   التواصل الاجتماعي لطلاب قسم اللغة الإنجليزية، جامعة قسنطينة

</div>

* Manel Mouleme

Faculty of Letters and Languages Department of Foreign Languages
University of Mentouri Constantine I (Algeria), linguisticsumc@gmail.com

Abstract: The current study seeks to shed light on the codification of Algerian Arabic. To achieve this goal, qualitative research based on corpus analysis has been employed. A total of twenty Facebook messages from English Department students have been examined. According to the findings, Latin letters have been used to portray Algerian Arabic. As a result, an Algerian Arabic sounds, letters, and symbols table has been created. The main contribution of this study is that it serves as a first step towards the codification of Algerian Arabic. As a result, Algerian Arabic keyboards should be included in mobile phones and laptops for better communication.

Keywords: Algerian Arabic, Modern Standard Arabic, orthography, social media.

<div dir="rtl">

ملخص: تسعى الدراسة الحالية إلى تسليط الضوء على تدوين اللغة العربية الجزائرية. لتحقيق هذا الهدف، تم استخدام دراسة نوعية تستند إلى تحليل المحادثة تم فحص ما مجموعه عشرين رسالة   على الفيس بوك من طلاب قسم اللغة الإنجليزية. وفقًا للنتائج، تم استخدام الحروف اللاتينية لتصوير اللغة العربية الجزائرية. نتيجة لذلك، تم إنشاء طاولة أصوات وحروف ورموز عربية جزائرية. وتتمثل المساهمة الرئيسية لهذه الدراسة في أنها تشكل خطوة أولى نحو تدوين اللغة العربية الجزائرية. نتيجة لذلك، يجب تضمين لوحات المفاتيح العربية الجزائرية في الهواتف المحمولة وأجهزة الكمبيوتر المحمولة لتحسين الاتصال

كلمات مفتاحية: العربية الفصحى، اللهجة الجزائرية العامية، الأبجدية العربية، التواصل الاجتماعي

</div>

Corresponding author: Manel Mouleme, e-mail: linguisticsumc@mail.com

1. INTRODUCTION

Algeria became a diglossic country par excellence due to the presence of Arabic. Modern Standard Arabic is the high variety used in official and formal situations, while Algerian Arabic is the low variety used in informal situations and everyday communication. Modern Standard Arabic (MSA) is standardised because it has grammatical rules and is taught in a formal setting; it also has orthography, whereas Algerian Arabic (AA) is not standardised because it is acquired as a mother tongue and has no writing systems. For a long time, Algerians used to communicate, whether officially or unofficially, using MSA. Yet, since the advent of the internet and the widespread use of social media, people have started using AA in their daily messages, whether on social media like Facebook and Instagram or in phone messages. Since AA is not codified, Algerians rely most of the time on French and Arabic keyboards to record their messages. However, when using French keyboards, some sounds are unique to Arabic. To code these sounds, Algerians use Latin letters or numbers. Therefore, the present study aims to investigate and describe the writing systems used to codify Algerian Arabic. To accomplish this goal, twenty (20) Facebook messages from English department students are being examined. Only seven messages, as

examples, are presented in this paper. Furthermore, the present study seeks to answer the following research questions: (1) Do university students use AA in their everyday communication? (2) On which language do students rely on when writing their Algerian Arabic conversations? (3) What is the orthography used by university students? To answer these questions and to reach the already stated aim, a sample of seven (7) messages is presented in this paper to show instances of how AA is written and codified using letters and numbers, and a list of all AA sounds with their corresponding letters is collected.

2. Literature Review

2.1 Related Works

Arabic dialects were not examined extensively in the past. They have recently sparked an interest. Some studies are interested in studying Arabic dialects, such as Dendane's (2015), in which he discussed how the fast transition from the mother tongue to the school language affects the quality of Algerian children's formal learning and the problems they confront in primary school. Further, Belfarhi (2019) investigated Algerian Arabic. In her paper, she explored the integration of independent elements from independent languages into one code, with written French integrated with spoken Algerian Arabic, creating an integrative code that is meaningful to Algerian readers and not French.

CODA (Conventional Orthography of Dialectal Arabic), which Habash made in 2012, is a standard way to write dialectal Arabic. Its main purpose is to help make computer models of Arabic dialects. Habash et al. (2012) wrote a paper in which they talked about CODA's design ideas and gave a full overview of its criteria as they applied to Egyptian Arabic. Later, Zribi et al. (2014) broadened the CODA rules to include the Tunisian dialect, and Jarrar et al. (2014) applied them to the Palestinian dialect. Furthermore, CODA authors from Egypt and Tunisia support the extension of CODA to other Arabic dialects to generate linguistic resources and computational models of Arabic dialects.

After developing conventional orthography for Egyptian and Tunisian Arabic dialects, Saadane and Habash (2015) attempted to do the same for Algerian Arabic. They highlighted the design elements of the Algerian CODA and provided a comprehensive explanation of its principles. At this juncture, Harrat et al. (2016) investigated the Arabic Algerian dialect, a non-resourced language for which no recognised resource is available to date. They presented the results of their first linguistic research, which introduced its most essential aspects and described the resources that they had generated from scratch for this dialect. Their work focused on the dialect spoken in Algiers and its surroundings. They showed all of the features associated with it, and then they introduced resources that they produced from scratch. This represented a cleaned corpus of the Algiers dialect aligned to MSA.

Baya and Kerras (2016) conducted another study on Algerian Arabic. In their article, they examined the role of the Algerian language on the social networking site Facebook. The purpose of this research is to compare the frequency of use of Algerian Arabic in the texts examined to that of Modern Standard Arabic. They discovered that spelling and style were not good due to the speed of writing on social networking sites. Moreover, writing is done with a Latin keyboard because, at Algerian universities, foreign languages are used more than Arabic.

2.2 Modern Standard Arabic: Arabic is an umbrella term used to cover all varieties of Arabic spoken in different countries in the Arab world. It is used from Oman in the east to Mauritania in the west, and it is the official language of 21 countries in the Middle East and North Africa (Procha´zka, 2006). It is a Semitic language and a member of the Afroasiatic language family (BELNAP, 2005).

MSA is nearly identical to classical Arabic (CA)[1] in phonology, morphology, and syntax, but it differs significantly in lexicon, phraseology, and style (Procha´ zka, 2006). MSA is significant because, as the only accepted medium of written and formal oral communication, it serves as the linguistic link that joins the Arab world together. Nevertheless, MSA must be learned in school because every Arabic speaker's native tongue remains his or her local dialect that is used in daily situations by all layers of society. As a consequence, MSA is almost entirely restricted to written communication and highly formal speech (news, official speeches, and academic discourse) (Procha´ zka, 2006).

The Arabic alphabet evolved from the Nabataean alphabet. Arabic, unlike Nabataean, is a cursive script. It is made up of 28 letters and is written from left to right. (BELNAP, 2005). The dominance of consonants in the Arabic writing system reflects a distinguishing feature of Semitic languages known as the root and pattern system. The majority of words are built around a three-consonant root. Arabic excels at utilising this morphological resource more than any other Semitic language. Each root is associated with a central concept. Semitic derivational morphology is founded on the systematic manipulation of these roots (BELNAP, 2005).

MSA is declared a national and official language and is used in different domains in Algeria. In education, it is used in all cycles: primary, middle, secondary, and even higher education. It is used as the language of instruction and is used in textbooks and lessons.

2.3 Algerian Arabic: It is also known as Derdja, vernacular Arabic, and Arabic dialect. It is the low variety of Arabic used in everyday communication at home, on the street, and for non-official talks. It is not standardised, and it is learned as a first language during childhood. Algerian Arabic has no grammar books or dictionaries. Furthermore, it has no literary heritage, and all poetry and novels are written in Modern Standard Arabic. It also has no orthography, so it has no written form. Nonetheless, since the advent of the internet, Arabic dialects have become increasingly popular in social networks and forums. Because of the ease with which people with low levels of education can communicate in dialect, it emerges on the web as a real communication language (Harrat et al., 2016). This dialect contains a lot of borrowed words from French. It has numerous variations that are primarily the result of both phases of Arabization and a long history of colonisation (Bougrine, Cherroun, & Ziadi, 2018). For instance, the Algerian word /vista/ is borrowed from French 'la veste'.

2.4 Orthography: Orthography is a human invention, along with the alphabet and script. Orthography is sometimes used interchangeably with spelling and the writing system**.** Instead of matching symbols to ideas or concepts, language structures are used. Writing systems are graphic ways to record language, and they can relate to speech in different ways based on different concepts: Words, morphemes, syllables, consonants, and vowels can all be represented using characters (Cooke, 2021). Based on written signs, writing systems enable readers to reconstruct linguistic messages. Orthographies are writing systems that are standardised in terms of (a) a collection of visual symbols (graphemes), such as signs, characters, and letters, as well as diacritical marks, punctuation marks, and so on; and (b) a collection of standards, such as rules for capitalization, punctuation, word boundaries, and orthographic usage (Coulmas, 2003, p. 35; see also Coulmas, 1996, 1380; Rogers, 2005, p. 2ff; as cited in Seifart, 2006).

Seifart (2006) defined orthography as a combination of a set of graphemes, like an alphabet, and a set of rules for how to use them. The third distinguishing feature is the standardisation and codification of

both the symbols and their usage. The actual visual shape of the graphemes used by a writing system, such as Latin or Arabic letters, is referred to as its script.

3. Methodology

3.1 Research Design

The study is qualitative in nature and based on a descriptive research method. A corpus analysis is used where twenty (20) university students' Facebook messages are analysed. The analysis attempts to answer the research questions and reach the already stated aims. It seeks to investigate how Algerian university students write Algerian Arabic in their Facebook messages.

Students from the English department at Fréres Mentouri University in Constantine make up the sample. They are asked to screenshot their daily conversations and send them to the researcher. Students are supposed to send new and old messages. Only twenty messages have been selected randomly to be analysed. Only seven messages are going to be presented as they are, and an English translation is included. Moreover, the phonetic transcription of the sounds is taken into account for comparison.

3.2 Data Analysis

Students' Samples

Here are some examples of conversations between students on Messenger in which they used Algerian Arabic written in Latin letters. These short conversations are translated into English.

Conversation 1

B: wach tdiri bach tsajli fl American corner
What have you done to enrol in the American Corner?
Normalement photocopie ta la carte w 2 photos
Normally a photocopy of your card and two photos
A: **Nti msajla fih**
Are you enrolled?
B: **wlh mani msjla**
I swear I am not enrolled
Jibi mla carte d ʿétudiant
So, Bring with you your student's card
W gulha n9ra anglais b3d
And tell her that I study English
Khli ghadwa nshlh w nfrzouha
Let it for tomorrow to clarify things
A**: Nchlh ya rabi**
See you then
Bn8
Good night
B**: good night**

The above conversation is between two students who want to enrol in the American Corner. The language used is Algerian Arabic, with some French words. Students are using Latin letters to write Algerian sounds. For instance, the sounds /ʃ/, /s/, /t/, /d/, /w/, /f/ and /l/ are represented using the Latin letters as follows: **ch**, **s**, **t**, **d**, **w**, **f**, and **l**. Moreover, the sound /ʒ/ is represented by **j.** Also, the letter **r** is used to represent the sound **/r/.**

The student used the letter a to represent the sound / ʕ /, as in the example 'ta,' which means 'of.' The letter **g** is used to represent the sound /g/, which does not exist in MSA. So, the student used the Latin letter, which already exists in French. However, the sound /q/ is represented by the number 9, like in the example **n9ra,** which means 'I study'.

Conversation2

A**: Lt ! Ling cl**
  Literary texts! Linguistics, culture
B**: ouiii**
 yes
 A**: T9dry tjbli gdwa bh ndyr photocopie**
  Can you brin git tomorrow to make photocopy
B**: Owkey**
ok
A**: dac 1000 mercii**
 1000 thanks
B**: bla mziyaaa Dinaaa**
  Never mind
Ni mlgithumch
I did not find it
Chfy tra nouhaaa
Ask also Nouha
Ey dert 3lya photocopie déjà
She has already made photocopy

The second conversation is between two students talking about their lectures. Here again, the different Arabic sounds are represented by the Latin letters. For instance, the sound /h/ is represented by the letter **h** in the example **mlgithumch,** which means **'*I did not find it'. The number 9 represents the sound /q/ in the word "T9dry," which means "you can" in English.* Moreover, the number 3 is also used in Algerian orthography. It is used for the sound / ʕ /, as in the example **3lya,** which means **at**.

Conversation 3

A**: mais, makamlch hadouk**
  But he did not finish
Ta3 math Tamarin
Excercices of mathematics
B**: normal ani 3labali hih Goulou bark yjibhm m3eh**
 Yes I know just tell him to bring them with him
B**: dcr nchlh**
Ok, God's willing

In the above conversation, the students were talking about the BAC exam. The sounds of AA in this conversation are represented by the same letters and numbers used in the above conversations, such as **b, ch, k, l, h, d, t, j.** However, the letters **i**, **a**, and **ou** represent the sound vowels /i /, /a /, and /u /, respectively, while the letter e represents the sound / ə/.

Conversation 4

> A**: ntmna ntlgew chi nhaar Twhchtk**
> *I wish to meet you one day I missed you*
> **Take me back to those days**
> B**: Mm ena koun ghir nwliw 9 ans**
> *Even me, I wish to go back 9 years ago*
> **Yn9solna 10 ans mn 3mrna**
> *Reduce 10 years from our age*
> **Yrdouna lel primaire**
> *Take us to primary school*
> A**: wlh ghir mch mliha tkbrii**
> *It is not good to get older*
> **Kulchi yzid 3lik**
> *Everything ….*

In this conversation, students missed each other. They wanted to be back at the primary school. This conversation is written using Latin letters and numbers, too. The sounds /t/, /h/, /m/, /n/, /l/, /k/, /r/, /d/, /ʃ/, /b/, /w/ and /sˤ/ are represented by using the Latin letters as follows: **t**, **h**, **m**, **n**, **l**, **k**, **r**, **d**, **ch**, **b**, **w**, and **s**. Moreover, the sounds /q/ and /ʕ/ are represented by the numbers **9** and **3,** respectively**.** In addition, the sound /ɣ/ is represented by the letters **gh.** However, the letters used to represent the sounds of AA are the same as those used in the previous conversations**.** As for the vowels, the sound /u/ is represented by the letter **u,** as in the example '**kulchi**', which means everything*. The letter **i** represents the vowel /i/.*

Conversation 5

> A**: Okay thanks**
> B**: ashl nuum f dnya**
> *The easisest number in the world*
> **A partir de semedi nchlh rani nkoun f 03**
> *From Saturday I will be in university 03*
> **Golii 9bl brk psq mindek nkoun n9ra**
> *Just tell me before because I study from time to time*
> **T7bii njiblk les couleurs o tkhyrii Thm**
> *Do you want to bring with me colors and you choose there?*
> A**: dcr ni ngulk 9bl mnji w n3ayyatlk tani**
> *I will tell you before I come and I will also call you*
> B**: dcr jibhum**
> *Ok bring them*
> **U're so nice**

This conversation is between a student and a seller. They used Latin letters and numbers to record their words. The majority of the AA sounds in previous conversations have been represented by letters. The sound /g/ is represented by the letter g. The sound **y** is represented using y or double **i** (**ii**), as in the

word '**dnya**', which means life, and the word '**tkhyrii**', which means you choose. Moreover, the sound /tˤ/ is represented by the letter t, as in the example n**3yytlk,** which means 'I will call you'. Furthermore, the sound /ħ/ is represented by the number 7, as in the word **t7bii,** which means 'do you like'

Conversation 6

A: Njiibhmlk ghdwa sbah 9bl matbdey t9ray tssa3dk?

Do you want me to bring them to you tomorrow before you start your lectures?

B: Yee. Lkan tsa3dk nta tani mch muchkul

Yes, if it suits you too, I have no problem

A: ouii nrml ana

Yes

Hya 3lblk kount dayr fi balii mn9rach ghdwa l3chyaa

I thought I do not have any lectures tomorrow evening

Ay darouna wa7d tkhlatt tssma t7tmt nro7 3labiiha mng3dch 3la 2:30

They made a kind of mess that's why I have to go and I will be buzzy at 2:30

Lsl ghdwa sba7 nchlh njibhoumlk l 01 hka

So, tomorrow morning I will bring them to you at 1, right?

B: Okaay assidi deal!! W bgeddah twelli m3a livraison!

Ok, it is deal! And how much does it coast with the delivery?

Tji 3la 9:30

Will you come at 9:30?

A: aha njibhmlk mn 3ndii

   You do not have to pay the delivery

Most of the AA sounds in this conversation are represented by Latin letters and numbers. The sound /ʒ/, for instance, is represented by the letter **j** as in the word '**njiibhmlk**', which means 'I will bring them to you'. In addition, the sound /tˤ/ is represented by a double **t** as in the word '**tkhlatt**', which means 'mess'. To stress syllables, the stressed one is marked by double letters, as it is in the word **tssma,** which means 'which means', and '**begeddah**', which means 'how much'.

Conversation 7

A: ha nsa9sik?

Can I ask you a question?

B: ouii tfaddal

Yes, you are welcome

A: khok y3agab bac sna?

Will your brother pass the BAC exam?

B: Aaaa nn ma3awedch youcef

No, youcef will not pass the BAC exam.

Habbass

 He gave up

A:Aaa dcr

ok

Wld khti hay3agab sna gali wach lmou9tarahat ta3 filou w ana ma3labalich

My cousin will pass the BAC exam this year and he has asked me about the suggestions of philosophy and I do not know**.**

B: att nsa9silk amie ta3i t3aggab Lkan 3labalha

I will ask my friend she is going to pass the exam and I will tell you

A: oki

ok

B : ki tripondili tofla ni ngulk

As soon as she answers me, I will tell you.

A: merci fati

Thank you fati

This conversation is about the BAC exam. They talked about philosophy exam questions. The sound /dˤ/ is represented using a double **d** (dd), as in the example **tfaddal,** which means 'are welcome'. The letter **kh** is used to record the sound / χ/ as in the word '**khok**', which means 'your brother'. Moreover, the sound /f/ is represented by **f**. The sound /p/ is represented by the letter **p.** There is another abbreviation used in this conversation: **att** for the word **attends**, which means 'wait'.

3.3 Discussion of Results

The present research paper aims to answer the following research questions: (1) Do university students use AA in their everyday communication? (2) On which language do students rely on when writing their Algerian Arabic conversations? (3) What is the orthography used by university students?

The examination of the students' messages shows that students send messages to each other at different moments (in the morning, afternoon, or evening) to talk about various topics such as cooking recipes, their studies, and so forth. When they write their messages, they use AA, and they may switch to French or English. They switch to French because almost all Algerians use it in their everyday communication. They also switch to English since this is their discipline. In the same line of thought, Sid-Houès found in his study (2008) that university students switch between AA and French in their conversations.

Students use AA words when writing their messages. Since AA has no written system, university students developed an AA orthography, which is composed of some Latin letters and numbers. Sounds that already exist in French and English have their corresponding letters such as /b/, /d/, /f/, /g/, /h/, /m/, /n/, /p/, /r/, /s/, /t/, /v/, /w/, and /z/. These sounds are represented by the following letters: **b, d, f, g, h, m, n, p, r, s, t, v, w,** and **z.**

Sounds that do not exist in French or English are represented by some letters and numbers. For instance, the letters **kh** stand for the sound / χ/. The letter t is doubled to represent the sound /t/. Further, the sound /ħ/ is recorded using the letter **h** or the number **7**. The letters **gh** stand for the sound/γ/, while the sound /sˤ/ is represented by the letter **s**. In this case, the letter s stands for the sounds /s/ and /sˤ/. Furthermore, the sounds /q/ and / ʕ / are represented by the numbers **9** and **3**, respectively.

As far as the vowels are concerned, the sounds /a/, /i/, and /u/ are represented by the letters **a**, **i**, **u,** or **ou,** respectively. The letter **e** is rarely used by students to represent the sound / ə /.

The results of this study are exactly the same as what Baya and Kerras (2016) found in their study. In their research, they looked at how AA is used on the social networking site Facebook. They found that writing is done with a Latin keyboard because, at Algerian universities, foreign languages are used more than Arabic. They thought that Algerian Arabic was a new language that needed grammatical research to build its grammar and write it down, just like vulgar languages had been done. At the end of their work,

they said that the Algerian language needs to be written down so that it can be recognised as its language and set apart from the Arabic world.

Habash (2012) created CODA (Conventional Orthography of Dialectal Ar-abic) as a standardised orthography for dialectal Arabic, with a focus on the creation of computer models of Arabic linguistic varieties. In the same line of thought, the present study has attempted to develop a written system of Algerian Arabic to facilitate communication via computers and mobile phones. The following table shows the different Latin letters and numbers that can be used to write Algerian Arabic:

Consonants

MSA sounds and letters	AA sounds (phoneme)	Letters and numbers used to represent the AA sounds	Examples taken from students' messages
أ /ʔ/	/ʔ/	a	**Ana njibhmlk mn 3ndii** *I will bring them to you*
ب /b/	/b/	b	**Am ybi3lk les timbres** *He sells the stamp to you*
ت /t/	/t/	t	**Muhim lktab brk myhamch chakhsiya** *What is important is the book, not the character*
ج /dʒ/	/d/	j	**Tji 3la 9:30** *She will come at 9:30*
خ / χ/	/ χ/	kh	**Khok y3agab bac sna?** *Does your brother pass the BAC exam this year?*
ح /ħ/	/ħ/	7 / h	**T7bii yjiblk les couleurs o tkhyrii** *Do you want him to bring the colors and you then choose*
د /d/	/d/	d	**Ana dert diglossia** *I have prepared diglossia*
ر /r/	/r/	r	**Ma3rf** *I do not know*
ز /z/	/z/	z	**Bla mzziya** *Never mind*
س /s/	/s/	s	**Hada nefssou likount habba nchrih** *It is the same one I wanted to buy*
ص /sˤ/	/sˤ/	s	**Ew sob talj f jbal lwahch** *The snow fell at djebel wahch mountain*
ش / ʃ/		ch	**Chkun kifi** *Who is like me?*
ض /dˤ/	/dˤ/	d	**Drk nchuf** *I will see*
ط /tˤ/	/tˤ/	t tt	**Koul wahd fikom 3ta une idée** *Everyone has a given idea*
ع / ʕ /	/ ʕ /	3	**Wlh ma3ef** *I swear I do not know*
غ /ɣ/	/ɣ/	gh	**Welit nhuwes ghir bdfa** *I want only to feel warm*
ف /f/	/f/	f	**Wellit nhuwes ghir ndfa** *I want only to feel warm*
ق /q/	/q/	9	**Raho wa9thum** *It is high time*
ك /k/	/k/	k	**Kont ra7a nsa9sik 3la lg** *I was about to ask you about linguistics*
ل /l/	/l/	l	**Lyoum win khadmet** *Today I started working*
م /m/	/m/	m	**Ntuma m3likumch?** *How are you?*
ه /h/	/h/	h	**Wchmn ktab gal fih hadi**

392

			In which book he said this
ن /n/	/n/	n	**Wch hwalha nana** *How is my grandmother?*
و /w/	w	w / oi / o	**Wzebda?** *And butter*
ي /j/	j/	y / ii	**Wynk Amira** *Where are you, Amira?*
---------	/g/	g	**Golili gdh ndir dgig w zebda?** *Tell me how much flour and butter shall I use?*
--------	/p/	p	**Ki tripondili tofla ni ngolulk** *Once the girl answers me, I will tell you*
--------	/v/	v	**3ndek vest ykhi?** *You have a jacket, haven't you?*

Vowels

ó	/a/	a e	**Ma3rf** I do not know
ó	/u/	o / u / ou	**Chufi rohi ngoulek** *Look go tell you*
ọ	/i/	I	**Hih** *yes*

4. CONCLUSION

This paper sums up a first attempt to work on Algerian Arabic orthography. Its main contribution is that it aims to develop a writing system for Algerian Arabic. This dialect is not codified, but because of the constant need to communicate in Facebook messages, university students have started to use AA. They have developed their writing system. All AA sounds are represented using Latin letters and numbers. There is a kind of social agreement among university students on how AA is codified.

However, when conducting this study, one of the limitations faced was the small sample size. Only seven conversations have been analysed. It is challenging to conduct research with small samples, and this is one of the disadvantages of qualitative research in comparison to quantitative research, which requires larger samples to ensure a representative distribution of the population.

Based on the findings of this study, it is recommended that all phones and laptops include an external keyboard. Moreover, AA dictionaries and books can be produced to help foreigners (non-speakers of Arabic) learn Algerian Arabic. It is also recommended for future research to examine the codification of Algerian Arabic using Arabic letters.

5. Bibliography List :

Baya, M. L. & Kerras, N. (2016). A Sociolinguistic Study of the Algerian Language. *Arab World English Journal,* 3, 141-154.

Belfardi, K. (2019). Languages' Integration in Algeria: Dialectal Text and French Graphic. *Open Journal for Studies in Linguistics, 2* (1), 51-58. https://doi.org/10.32591/coas.ojsl.0202.02051b

Belnap, R. (2005). Arabic. In P. Strazny (Ed.), *Encyclopedia of Linguistics* (Vol. 1, pp. 74-77). New York: Fitzroy Dearborn.

Bougrine, S., Cherroun, H., & Ziadi, D. (2018). Prosody-based Spoken Algerian Arabic Dialect Identification. *Procedia Computer Science*, 128, 9-17. https://doi.org/10.1016/j.procs.2018.03.002

Cooke, G. (2021). Orthography. In J. Stanlaw (Ed.), *The International Encyclopedia of Linguistic Anthropology.* (pp. 1-6). Hoboken, NJ: Wiley-Blackwell.

Dendane, Z. (2015). Dialectal Arabic, MSA and Formal Education in Algeria. *Revue Traduction et Langues, 14*(1), 100-107.

Habash, N., Diab, M., & Rambow, O. (2012). Conventional Orthography for Dialectal Arabic. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12),* 711-718. Istanbul, Turkey: European Language Resources Association (ELRA)

Harrat, S., Meftouh, K., Abbas, M., & Hidouci, W. (2016). An Algerian dialect: Study and Resources. *International journal of advanced computer science*, 7 (*3*), 384-396.

Jarrar, M., Habash, N., Akra, D., & Zalmout, N. (2014). Building a Corpus for ePalestinian Arabic: a preliminary study. *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP)*, 18-27. Doha, Qatar: Association for Computational Linguistics. http://dx.doi.org/10.3115/v1/W14-3603

Prochāzka, S. (2006). Arabic. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (pp. 423-431). Amsterdam: Elsevier Science.

Saadane, H. & Habash, N. (2015). A Conventional Orthography for Algerian Arabic. *Proceedings of the Second Workshop on Arabic Language Processing*, 69-79. Beijing, China: Association for Computational Linguistics. http://dx.doi.org/10.18653/v1/W15-3208

Seifart, F. (2006). Orthography development. In J. Gippert, N.P. Himmelmann, & U. Mosel (eds.), *Essentials of Language Documentation* (pp. 275-299). Berlin: Mouton De Gruyter.

Sid-Haouès, A. (2008). *Code Variation among Algerian University Students*. Unpublished Doctoral Thesis. Constantine: Fréres Mentouri University.

Zbiri, I., Boujelbane, R., Masmoudi, A., Ellouze, M., Belguith, L., & Habash, N. (2014). Conventional Orthography for Tunisian Arabic. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 2355-2361. Reykjavik, Iceland: European Language Resources Association (ELRA)

---

[1] Classical Arabic is the Koranic language; it is considered the formal version used in the Arabic peninsula. It was also the pre-Islamic era's poetic language, as well as the language during Islamic history.