

# WIKIMEDIA AND ARABIC SOCIOLINGUISTICS

#### Houcemeddine Turki

Faculty of Medicine of Sfax, University of Sfax, Sfax, Tunisia

#### Anass Sedrati (Q&A)

National Institute of Posts and Telecommunications, Rabat, Morocco

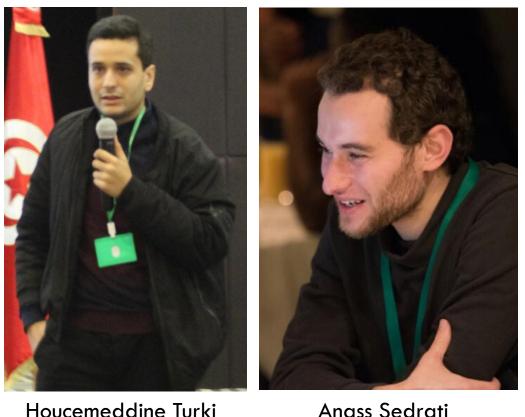








# **ABOUT US**



Houcemeddine Turki

#### Houcemeddine Turki (User:Csisc)

Medical student, University of Sfax (Pre-Med 2018)

Wikimedian since 2009 (IP editing since 2007)

Active in Wikipedia, Wikidata, Wikimedia Incubator, Metawiki, and Wikimedia Commons

Member, Wiki Project Med, Wikimedia R&D Tunisia, Wikimedia and Libraries User Group Steering Committee, WikiIndaba **Steering Committee** 

#### Anass Sedrati (User:Anass Sedrati)

Ph.D. Student in Telecommunications, National Institute of Posts and Telecommunications, Rabat, Morocco; B.A. MENA Cultures and International Relations, Stockholm University, Stockholm, Sweden

Co-Founder, Wikimedia MA User Group

Wikimedian since 2013

Active in Wikipedia, Wikidata, Wikimedia Incubator, Metawiki, and Wikimedia Commons

### **TEAM MEMBERS**



Rafik Zribi IESEG School of Management, France





Imed Adel University of Sousse, Tunisia





Reda Benkhadra Wikimedia MA User Group





**Ismael Zniber** Wikimedia MA User Group





Omar Tounsi-Sebbahi Wikimedia MA User Group



### **TEAM MEMBERS**



Mohammed Bachounda Algeria Wikimedians User Group





Reda Kerbouche Algeria Wikimedians User Group





Nacim Aboura Algeria Wikimedians User Group





Mekkia Bennama Algeria Wikimedians User Group



# LIAISONS

### Linguistics Liaison



Lameen Souag Université Sorbonne-Paris-Cité, France



### Library Liaison



Zeineb Takouti Diocesan Library of Tunis, Tunisia



### Wikidata Liaison



Helmi Hamdi Université de Sherbrooke, Canada



# LIAISONS

### **Computer Science Liaison**



Tariq Daouda Harvard University, United States



### Social Sciences Liaison



Nassim Regragui Copenhagen Business School, Denmark



## **RESEARCH ADVISORS**

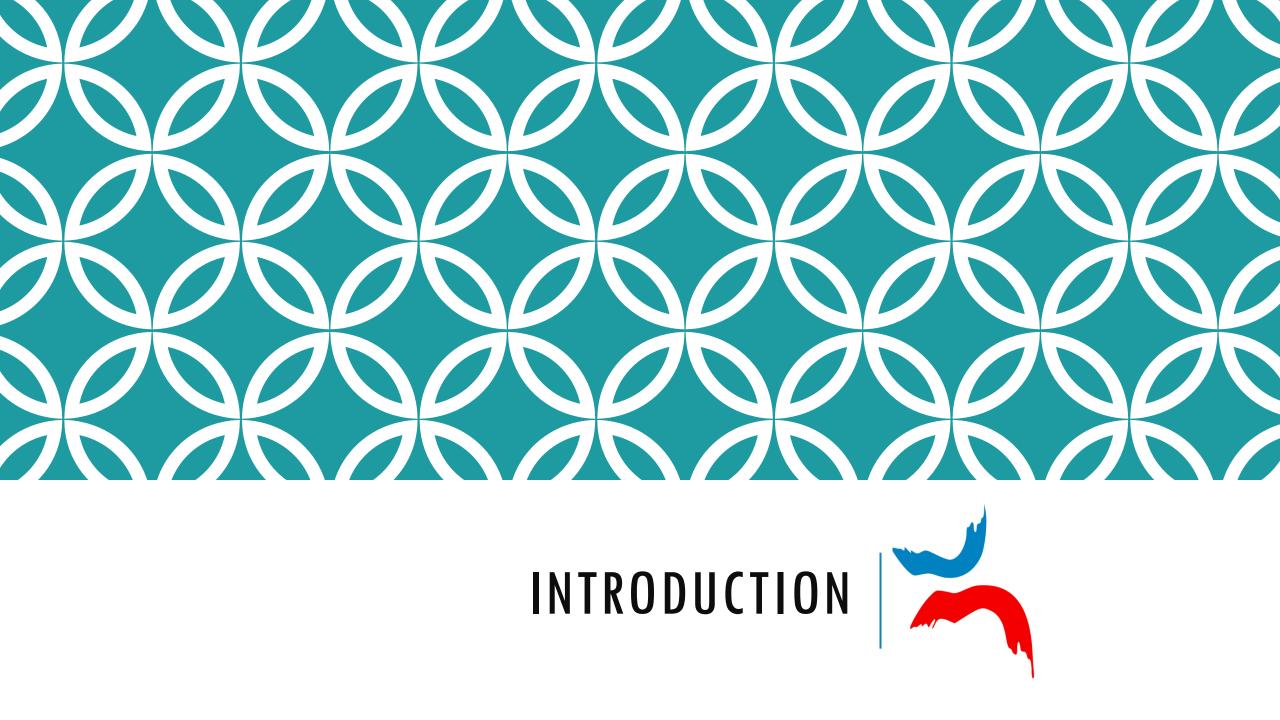
Mohamed Maamouri University of Pennsylvania, United States



Dominique Caubet Université Sorbonne-Paris-Cité, France





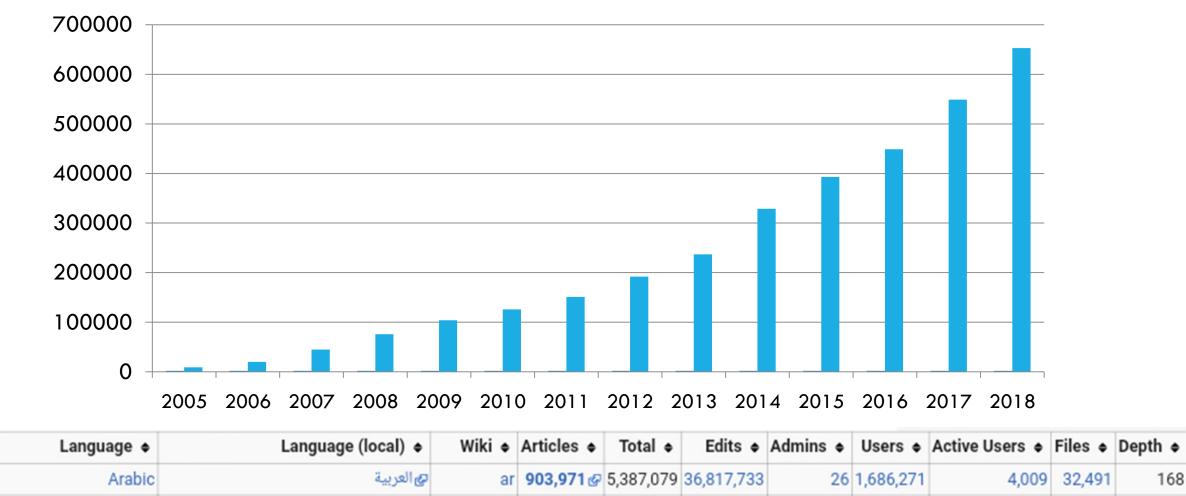


# **GROWTH OF ARABIC WIKIPEDIA**

Nº ♦

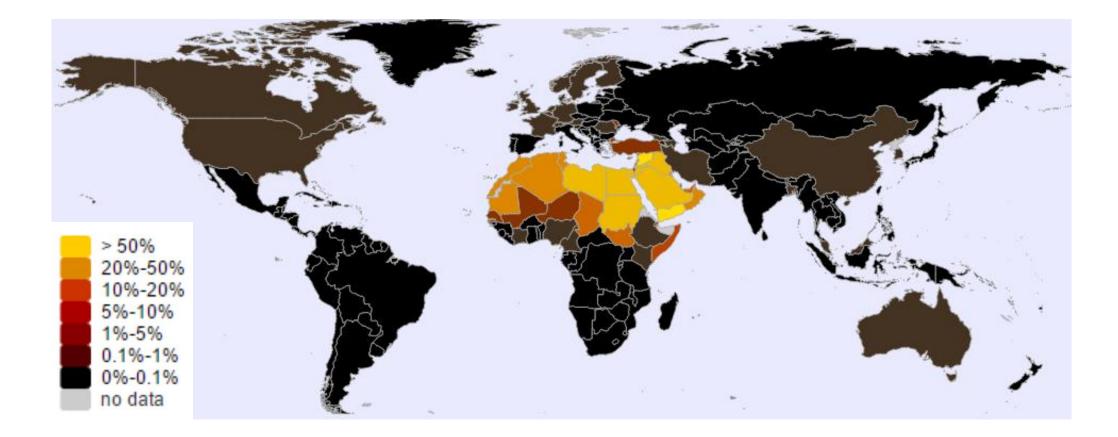
17

https://stats.wikimedia.org/FR/ChartsWikipediaAR.htm



### SHARE OF PAGEVIEWS TO ARABIC WIKIPEDIA

https://stats.wikimedia.org/wikimedia/animations/wivivi/wivivi.html



### **REASON: SOCIOLINGUISTICS** LANGUAGES OF THE YEAR 150 C.E.

https://commons.wikimedia.org/wiki/File:Roman\_provincial\_languages\_150CE.png



# **REASON: SOCIOLINGUISTICS**

When Classical Arabic had spread over Middle East and North Africa during the seventh century, it entered in contact with the languages already spoken in the region.

This caused the creation of adapted varieties of Arabic that are easier to use by each community. Each variety took into consideration several features of the local languages of each region. This let every dialect distinct from Classical Arabic and from other varieties of Arabic.

This situation became more critical over years because of the different contact between each Arabic dialect and foreign languages such as French, English, Italian and Turkish.

### **REASON: SOCIOLINGUISTICS** TWO INITIATIVES TO FIGHT THE LACK OF INTELLIGIBILITY OF CLASSICAL ARABIC TO ARAB PEOPLE

```
Modern Standard Arabic (19th century)
```

Replacing unintelligible lexical and grammatical structures with understandable structures from Eastern Arabic dialects

Maghrebi dialects not taken into consideration

```
Functional Arabic (1967)
```

Replacing unintellgible mainly technical words of Modern Standard Arabic by common Maghrebi Arabic words

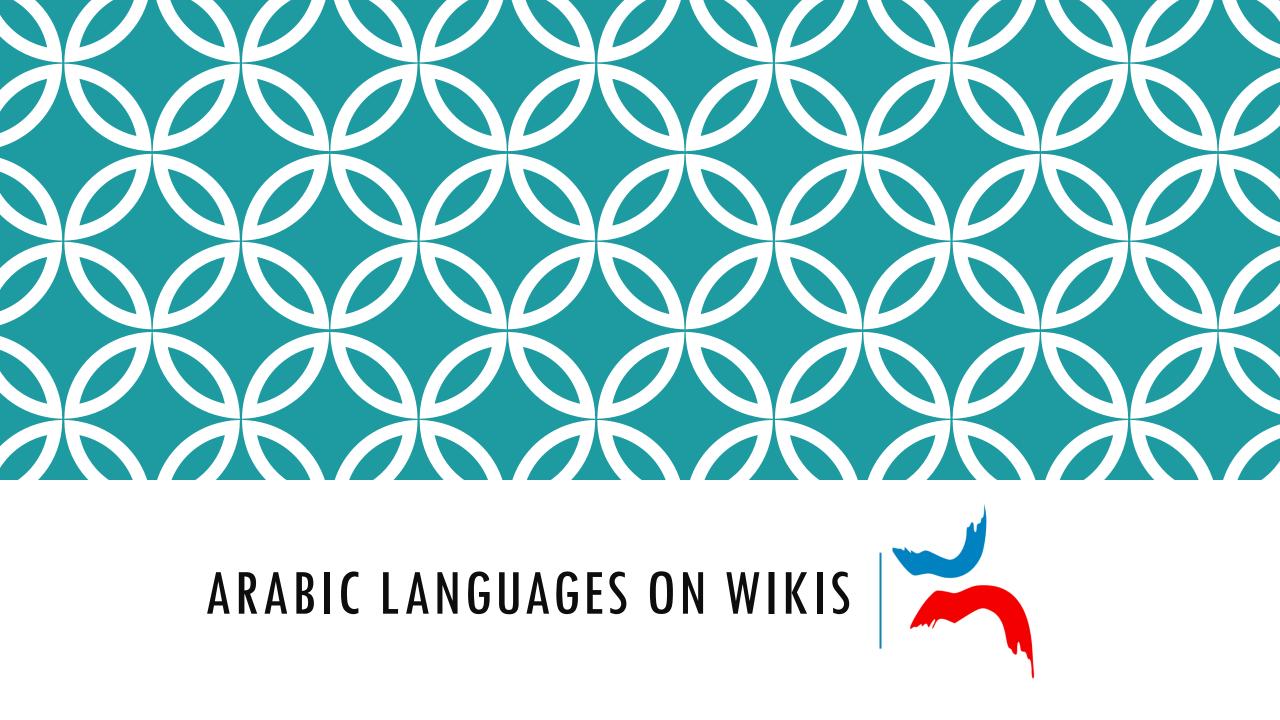
Not recognized by Middle Eastern speakers of Arabic

Consequences on Arabic Wikipedia

Lack of intelligibility of scientific articles on Wikipedia to Maghrebi users

Biased to Middle Eastern ideas and cultural thoughts

 $\rightarrow$  Adding support of Arabic languages to WMF wikis is needed to solve these matters



#### 24-Nov-Created

WIKIPEDIA EGYPTIAN: LIMITED OUTPUT, MANY Wikipedia Egyptian Arabic

# 2008

#### https://arz.wikipedia.org

#### N⁰ Language (local) Wiki Articles **Edits Admins Files Depth** Total Users Active Users Language arz 20,598 🖸 177,770 🗗 (Masri) مصری 114 Egyptian Arabic 873.455 6 113.260 133 1.457 286

#### ابادة الشعوب

**إبادة الشعوب[1]** هي تدمير جزئي أو كلى لشعب أو جنس أو جماعة دينية أو عقائدية عن طريق قتل حكومات لعدد من أفرادهم. الأمم المتحدة تعتبر هذا العمل جريمة دولية تستحق معاقبة مرتكبيها بطريقة دولية عن طريق محاكمتهم دولياً. من أشهر محاكمات جريمة إبادة الشعوب كانت محاكمات نورمبرج سنة 1945 التي حاكمت النازيين اللذين ارتكبوا جرائم إباده ضد اليهود و غيرهم [2][3][4]



#### مصادر [مصدر التعديل]

1.1 بیکتب انجلیزی کده:Genocide.

.Rubinstein, W.D. (2004). Genocide: a history. Pearson Education. p. 308. ISBN 978-0-582-50601-5 1.2 William Schabas. "Genocide in International Iaw: the crimes of crimes". Cambridge University Press, 2000, 1.3 p. 25: "Lemkin's interest in the subject dates to his days as a student at Lyoy University" ".followed attempts to prosecute the perpetration of the massacres of the Armenians

Charles H. Anderton, Jurgen Brauer, ed. (2016). Economic Aspects of Genocides, Other Mass Atrocities, 1.4 and Their Prevention. Oxford University Press. ISBN 978-0-19-937829-6

سے الصفحه دی فیها **تقاوی** مقاله عن **جریمه** و انت ممکن تساعد ویکیپیدیا مصری علشان تكبرهاي.



1938

الان تورينج

(تحويل من آلان تورينج)



STUBS, SEVERAL ARTICLES IN MSA

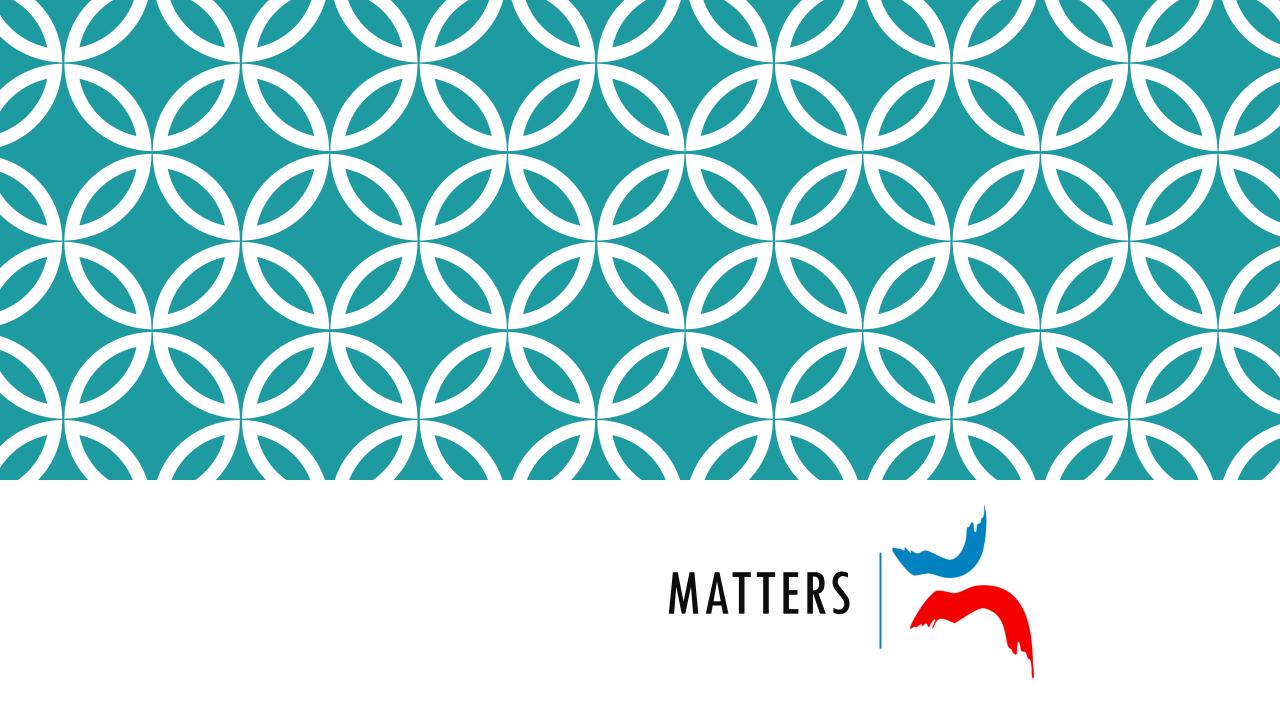
science

## OTHER WIKIS STILL IN WIKIMEDIA INCUBATOR

Project	Official status	Status since
Wikipedia Moroccan	Verified as eligible	06- Apr- 2008
Wikipedia Tunisian	Discussion	27- Sep- 2014
Wikipedia Hijazi Arabic	Discussion	24- Feb- 2016
Wikipedia Levantine Arabic	Discussion	14- Jun- 2015
Wikipedia Maghrebi Arabic	Discussion	30- May- 2015
Wikipedia Mesopotamian Arabic	<mark>On hold</mark>	30- Jul- 2018
Wikipedia Jazayri	Verified as eligible	04- Jul- 2009
Wiktionary Jazayri	Discussion	21- May- 2015

https://meta.m.wikimedia.org/wiki/Requests\_for\_new\_languages

Wikipedia Algero- Moroccan Arabic	Rejected	23- Apr- 2007	
Wikipedia Bahrani Arabic	Rejected	27- Jul- 2018	
Wikipedia Libyan Arabic	Rejected	22- Mar- 2018	
Wikipedia Northwest African Arabic	Rejected	22- Mar- 2018	
Wikipedia South Levantine Arabic	Rejected	07- Nov- 2008	
Wikivoyage Egyptian Arabic	Rejected	06- Jun- 2019	
Wiktionary North Levantine Arabic	Rejected	25- Jan- 2019	
Wikipedia Sudanese Arabic	Rejected	17- Oct- 2009	



# LACK OF NORMALIZATION

Orthography guidelines for writing Arabic languages in Arabic script are limited and quite imprecise

Many Latin script-based writing systems for Arabic languages: Arabizi, DMG...

The rules of Latin Script-based orthography are not well defined

Arabizi is widely used in social media. However, there is a lack of consensus of using it for formal purposes

# THOUGHTS ABOUT ARABIC LANGUAGES

Initiatives to add support of regional Arabic languages to WMF wikis are always offended by pro-MSA Arabic Wikipedia community

- The Arabic community sees the promotion of the formal use of regional Arabic languages as well as the use of Latin Script to transcribe them as a long-term project of the occident to destruct Modern Standard Arabic, Nationalist ideas and Arab cultural and religious heritage
- They use several politically influenced projects for the promotion of regional Arabic languages such as Said Akl project as examples
- They always tend to underestimate regional Arabic languages and consider them as accents of Arabic language although this is absolutely not accurate
- They always show that Modern Standard Arabic is the unique language of the Arab world although most of the Arab people rarely or never use it in real life

^ Discussion

- **Oppose** --Alaa :)..! 11:40, 25 November 2017 (UTC)
  - @ملاء العلام: Ehh why? --Liuxinyu970226 (talk) 11:43, 25 November 2017 (UTC) @Liuxinyu970226: Its not a language! Tunisian is not a separate language from Arabic, its just an accent. The differences between it and other Arabic accents are not extreme to the degree that makes it a separate language --Alaa :)..! 11:47, 25 November 2017 (UTC)
- Oppose for the same reason that alaa provided-- مصعب (talk) 12:20, 25 November 2017 (UTC)
- Oppose What a absurd coincidences!
  Migration season to the absurd. -- صالح (talk) 19:54, 5 July 2019 (UTC)

# MOST OF THESE THOUGHTS ARE NOT ACCURATE

Regional Arabic languages were first used in written form by Islamic theologists and mystic poets since the 15th century

It is absolutely accurate that occidental colonizers were based on regional Arabic languages to communicate with the colonized Arab societies. However, they also found Modern Standard Arabic useful to declare decisions and statements related to all colonized Arab states.

During the Arab Renaissance, Arab scholars used regional Arabic languages to write several important literary masterpieces and research publications so that they can reach and cultivate mostly illiterate Arab people.

Nationalist politicians such as Neo-Destour members in Tunisia studied Arabic dialects and then used them to communicate with people and raise their awareness about Patriotic issues



Category:Dictionnaire arabe-français (1938)

Une page de Wikimedia Commons, la médiathèque libre.

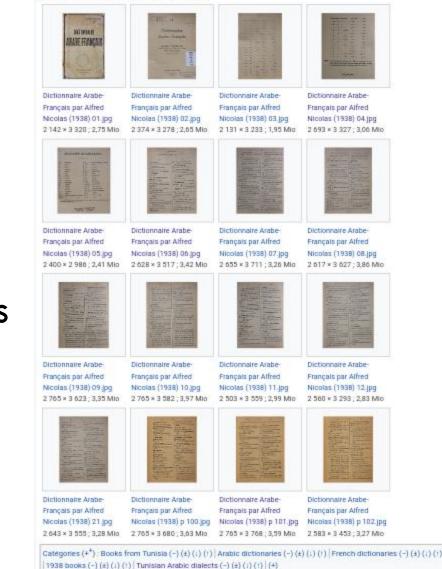


**n** 1

Média dans la catégorie « Dictionnaire arabe-français (1938) »

Cette catégorie comprend 340 fichiers, dont les 200 ci-dessous.

#### (page précédente) (page suivante)



# **REFERENCE BOOKS ON COMMONS**

Scanned edition of more than fifteen public domain reference books about Arabic Languages are uploaded to Commons to raise awareness about matters related to Arabic Sociolinguistics

### ADDING DETAILED LANGUAGE DESCRIPTIONS IN 🚨 Csisc 🏛

### THE ENGLISH WIKIPEDIA

#### Tunisian Arabic

Main page

Contents

Featured content

Current events

Random atticle

Wikipedia store

About Wikipedia

Recent changes

What links here

Related changes

Unicad file

Special pages

Permanent Ink

Wikidata Item

Cite this pape

In other projects

Wikivoyage

Print/export

Create a book

Download as PDR

Printable version

Languages

Aragonés

Asturianu

Brezhoneo Deutsch

Ελληνικά Español

فارمى

Francals Հայերեն

Hrvatski

Maiti

Nederlands 日本語

Plemontèls Português

Русский

Svenska Українська

اردو 中文

Bahasa Indonesi Italiano Македонска

العربية

Wikimedia Commons

Page information

Contact page

Tools

Community portal

Interaction

Help

Donate to Wikipedia

#### From Wikipedia, the free encyclopedia

This article is about a general description of the language. For the grammatical cases and conjugation, see Tunisian Arabic morphology

> This article or section should specify the language of its non-English content, using {{lang}}, with an appropriate ISO 639 code. See why. (May 2019)

Tunisian Arabic, or simply Tunisian, is a set of dialects of Maghrebi Arabic spoken in Tunisia.<sup>[5]</sup> It is known among its over 11 million speakers as: تونسى, romanized: Tounsi ['tu:nsi] ( listen),[6] "Tunisian"[7] or Derja "everyday language" to distinguish it from Modern Standard Arabic, the official language of Tunisia. As part of a dialect continuum, Tunisian merges into Algerian Arabic and Libyan Arabic at the borders of the country. Tunisian Arabic's morphology, syntax, pronunciation, and vocabulary are considerably different from Modern Standard Arabic or Classical Arabic. Like other Maghrebi dialects, it has a vocabulary that is mostly Arabic with a significant Berber, Latin<sup>[8][9]</sup> and possibly Neo-Punic<sup>[10][11]</sup> substratum. However, Tunisian has also many loanwords from French,[12] Turkish.<sup>[12]</sup> Italian<sup>[12]</sup> and the languages of Spain.<sup>[12]</sup>

Tunisian Arabic is mostly intelligible to speakers of other Maghrebi dialects but is hard to understand or is unintelligible for speakers of Middle Eastern Arabic.<sup>[8]</sup>

Contents [hide]

Multilingualism within Tunisia and in the Tunisian diaspora makes it common for Tunisians to code-switch, mixing Tunisian with French, English, Standard Arabic or other languages in daily speech.<sup>[13]</sup> Within some circles, Tunisian Arabic has thereby integrated new French and English words, notably in technical fields, or replaced old French and Italian loans with standard Arabic words.[13][14]

However, code-switching between Tunisian Arabic and modern standard Arabic is mainly done by more educated and upper-class people and has not negatively affected the use of more recent French and English loanwords in Tunisian.<sup>[13]</sup> Moreover, Tunisian Arabic is closely related to Maltese.[15] which is a separate

language that descended from Tunisian and Siculo-Arabic.[16][17] Maltese and Tunisian Arabic have about 30-40% spoken mutual intelligibility.<sup>[18]</sup>

Tunisian Arabic Touns! تونسى 9 ٥ \* പ Pronunciation (tuns) (0 actes) Tunisia, north-eastern Algeri Tunisians 11.2 million native (2014 census) Afro-Asiatio Semitic Central Semitic Arabic Manhrebi Arabio Tunisian Arabic Arabic script, Latin script Signed forms Tunisian Sign Language Official status As a variety of Maghrebi Arabi on 7 May 1999 (Not ratified due to several Constitutiona Matters) 2 Language codes aeb tuni1259/0 H

Native to

Ethnicity

speaker

Language

Writing

Recognised

language in

ISO 639-3

Glottolog



#### Talk Sandbox Preferences Beta Watchlist Contributions Log out

Read Edit View history 🏠 More 🗸 Search Wikipedia

Q

#### Tunisian Arabic morphology

From Wikipedia, the free encyclopedia

Article Talk

This article or section should specify the language of its non-English content, using {{lang}}, with an appropriate ISO 639 code. See why. (April 2019)

The grammar, the conjugaison and the morphology of Tunisian Arabic is very similar to that of other Maghrebi Arabic varieties.<sup>[1]</sup> It is based on Classical Arabic and influenced by Berber languages and Latin, with some morphological inventions. The Berber influence is more noticeable in Pre-Hilalian dialects.<sup>[1]</sup>

edia					
portal	Contents [hide]				
ges	1 Pronouns				
e	1.1 Personal pronouns				
	1.2 Possessive pronouns				
ere	1.3 Indirect object pronouns				
iges	1.4 Indefinite pronouns				
	2 Interrogative pronouns				
55	3 Articles				
ink.	3.1 Definite articles				
ation m	3.2 Demonstrative articles				
iii iii	3.3 Possessive articles				
~	4 Modal verbs				
	4.1 Hähü (To be, drawing attention to the presence of the referent)				
x	4.2 Råhū (To be, with more intensity by emphasizing it)				
s PDF	4.3 Mähü (To be, as an evidence marker or in a questioning manner as in tag questions)				
rsion	4.4 Q8sid (To be, at the immediate moment)				
0	4.5 Najjam (Could)				
~	4.6 Ynajjam (Can, To be able to)				
✓ Edit links	4.7 Haqū (Should)				
	4.8 Karū (Would better, stronger intensity than should)				
	4.9 Yilzmü (Have to)				
	4.10 Lāzmū (Must)				
	4.11 Mädäbih (Had better)				
	5 Discourse markers				
	5.1 Evidence markers				
	5.2 Conclusion markers				
	6 Preverbal markers				
	7 Verb conjugation				
	7.1 Perfective and Imperfective tenses				
	7.1.1 Regular verbs				
	7.1.2 Weak verbs				
	7.1.3 Irregular verbs				
	7.2 Future tense				
	7.3 Imperative tense				

1 Classification 2 History

# ADDING SUPPORT OF ARABIC LANGUAGES ON WIKIDATA

LABELS, DESCRIPTIONS AND ALIASES OF WIKIDATA ENTITIES

دينه وليا		۹۸ تونسې ۱	اللاحی تفضیلاتی تجریس لیستة الیامات (ای تجعیها سناهمات (امرچ)    اللاحی تفضیلاتی تجریس لیستة الیامات (ای تجعیها سناهمات (امرچ)    اللاحی تفضیلاتی (ای لاحی)    اللاحی (ای لاحی)    اللاحی (ای لاحی)
wikidata	(Q1315		
نوع من النيا	لمريية		
اباج الأولانية بوانة المبتنع القام المتروع إنشاء عصر جديد			
الشاء لكسيم جديدة البديلات الهند باج بالزهر Query Service		اصوفة <sup>العربية</sup> . - عرجع	
باللزب من هنا Random Primary Sources item معاونة البرع طبطن و تمريح اعمل كتاب			

#### ADDING SUPPORT OF ARABIC LANGUAGES ON WIKIDATA L-id Lexeme

#### LEXICOGRAPHICAL DATA

	Lexeme Dis	cu ssion	Lire Listedesibdles Plus v Rechard	er dan s Wikida ta 🛛 🍳
WIKIDATA	(L1)	ama	X	<b>∲</b> modifi <del>a</del>
Accuell Communaut é Bistro Créer un nouved d'ém ent Créer un nouvedu lexém e Modification srécert és Él ém ent au hasard Query Service À proximit é Bandom Frimary	Langu e sum Naturenom Déclar		mis-x-Q401	
	décritpar	8	Penn sylvania Sumerian Dictionary anglate + 1 référence	<b>∦</b> modtfier
Sourcesitem 🍐 😳				+ ajouter un e valeur
Faireun don Imprim er / exporter				+ ajouter un e déclaration
Télécharger en PDF Outlis	Sens			
Pag esil des Suivi d'espag esil des Pag es spéciales Lien permanient Information sur la pag e URI du concept Chier cettepagie Ancestors Primary Sources II d Importier Lesint erwikis Ajout automatique Vérifier Leilien BrowsePrimary Sources A back totop A		anglais français laŭn croate allemand arménien russe polonais albanais italien turc asturien basque ukrainien malayalam	mother mitre mater majka Mutter duujn Matb matka nithië matre anne ma ama ати ати сте222 щи	<b>₽</b> modifie

a meaning carrying part of a language, such as a word or a phrase

#### Lemma

one

many

many

standard form or dictionary form of the lexeme for verbs this is usually the infinitive form, for a noun the nominative singular, etc.

#### Lexical category

one also known as the part of speech or word class defines the lexeme to be either a noun, or a verb, or an adjective, etc. the set of possible values is open and taken from the Wikidata items

#### Language

chosen in an open list from Wikidata items

#### Statements

e.g. derived-from, region, period, homonym, etc.

#### Forms

one

many

many

specific, conjugated or inflexed forms of the lexeme

Representation the actual string value realizing a given form

#### Grammatical features

e.g. normative, comparative, past tense, etc.

#### Statements

e.g. region, period, pronunciation, etc.

#### Senses

Gloss short description, translatable in all languages of Wikidata

#### Statements

e.g. translations, synonyms, connotation, register, usage example, refers-to-concept, etc.

## WIKIS IN ARABIC LANGUAGES

Limiting the creation of wikis in Arabic Languages to two or three varieties of Arabic including Egyptian and Algerian so that we can have sufficient number of users for each newly created wiki

Each article should be limited to one paragraph so that it can be verified and adjusted by a small number of admins

Each wiki should use a lexicon that is common in the region it represents. For example, Wikipedia Algerian should use words that are also commonly used in Tunisia and Morocco

### **REFERENCES** INTRODUCTION

Visscher, M. S. (2012). Landscape of Languages – The position of provincial languages in the Roman Empire in the first three centuries AD. Leiden University.

Sayahi, L. (2014). Diglossia and language contact: Language variation and change in North Africa. Cambridge University Press.

Said, E. W. (2004). La langue arabe, la Rolls et la Volkswagen. Le Monde diplomatique, (8), 17.

Kaye, A. S. (1970). Modern Standard Arabic and the Colloquials. Lingua, 24, 374-391.

Mahmoud, Y. (1982). Towards a functional Arabic. Al-'Arabiyya, 15(1-2), 82-89.

Maamouri, M. (1998). Language Education and Human Development: Arabic Diglossia and Its Impact on the Quality of Education in the Arab Region. International Literacy Institute.

Sedrati, A., & Ait Ali, A. (2019). Moroccan Darija in Online Creation Communities: Example of Wikipedia. In VIIth International Congress on Moroccan Arabic (p. 26). Universidad Complutense Madrid.

### REFERENCES MATTERS

Said, E. W. (2004). La langue arabe, la Rolls et la Volkswagen. *Le Monde diplomatique*, (8), 17.

Sedrati, A., & Ait Ali, A. (2019). Moroccan Darija in Online Creation Communities: Example of Wikipedia. In VIIth International Congress on Moroccan Arabic (p. 26). Universidad Complutense Madrid.

Allehaiby, W. H. (2013). Arabizi: An Analysis of the Romanization of the Arabic Script from a Sociolinguistic Perspective. Arab World English Journal, 4(3), 52-62.

Abu Elhija, D. A. (2014). A new writing system? Developing orthographies for writing Arabic dialects in electronic media. Writing Systems Research, 6(2), 190-214.

Salameh, F. (2010). Language, memory, and identity in the Middle East: The case for Lebanon. Lexington Books

### REFERENCES MATTERS

Jouad, H. (2000). Le chant melhûn. Horizons Maghrébins-Le droit à la mémoire, 43(1), 32-45.

Fakhfakh, N. (2007). Le répertoire musical de la confrérie religieuse «Al-Karrâriyya» de Sfax (Tunisie). Université de Paris VIII.

Lecerf, J. (1932). Littérature dialectale et renaissance arabe moderne. Bulletin d'études orientales, 2(2), 179-258.

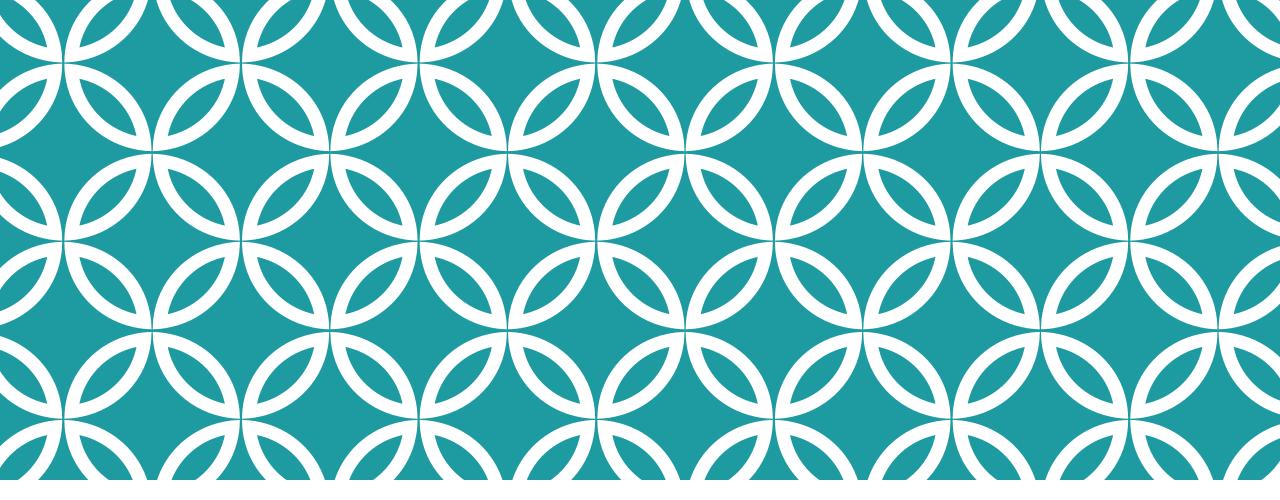
Beït Al-Hikma (2012). L'Encyclopédie Tunisienne. Académie Tunisienne des sciences, des lettres et des arts.

Brown, L. C. (2001). Bourguiba and Bourguibism revisited: reflections and interpretation. *The Middle East Journal*, 55(1), 43-57.

#### Linguistic works by Neo-Destour members:

Marçais, W., & Guiga, A. (1925). Textes arabes de Takroûna. Éditions E. Leroux.

Marçais, W., & Farès, J. (1933). Trois textes arabes d'El-Hamma de Gabès. Journal asiatique, 1-88.



Houcemeddine Turki User:Csisc turkiabdelwaheb@hotmail.fr

Anass Sedrati User:Anass Sedrati asedrati@wikimedia.org

