

Gulf Arabic Linguistic Resource Building for Sentiment Analysis

Wafia Adouane¹ and Richard Johansson²

{¹Department of Philosophy, Linguistics and Theory of Science,

²Department of Computer Science and Engineering}

University of Gothenburg

wafia.gu@gmail.com, richard.johansson@gu.se

Abstract

This paper deals with building linguistic resources for Gulf Arabic, one of the Arabic variations, for sentiment analysis task using machine learning. To our knowledge, no previous works were done for Gulf Arabic sentiment analysis despite the fact that it is present in different online platforms. Hence, the first challenge is the absence of annotated data and sentiment lexicons. To fill this gap, we created these two main linguistic resources. Then we conducted different experiments: use Naive Bayes classifier without any lexicon; add a sentiment lexicon designed basically for MSA; use only the compiled Gulf Arabic sentiment lexicon and finally use both MSA and Gulf Arabic sentiment lexicons. The Gulf Arabic lexicon gives a good improvement of the classifier accuracy (90.54 %) over a baseline that does not use the lexicon (82.81%), while the MSA lexicon causes the accuracy to drop to (76.83%). Moreover, mixing MSA and Gulf Arabic lexicons causes the accuracy to drop to (84.94%) compared to using only Gulf Arabic lexicon. This indicates that it is useless to use MSA resources to deal with Gulf Arabic due to the considerable differences and conflicting structures between these two languages.

Keywords: Modern Standard Arabic (MSA), Gulf Arabic, sentiment analysis, Arabic Natural Language Processing, Gulf Arabic sentiment lexicon

1. Introduction

Arabic sentiment analysis is an active area where many works have been done recently for various topics. However, most of these works are Modern Standard Arabic (MSA) based even though most Arab people use their own dialects to express themselves online. The major obstacle to face is the non-existence of linguistic resources for Gulf Arabic, namely lexicons for sense disambiguation, sentiment lexicon, annotated data to train on, Part of Speech (PoS) tagger and above all the complexity of this dialect itself. In this paper, we try to address sentiment analysis task for Gulf Arabic using machine learning. Therefore, we started with building linguistic resources for this dialect, i.e. annotated corpus and sentiment lexicon. It is worth mentioning that there are considerable differences between these two languages and using MSA sentiment lexicon to deal with Gulf Arabic for our task has been proven to be useless, as will be shown. This paper is organized as follows: first, we give an idea of the most recent works dealing with MSA-based sentiment analysis and some attempts to deal with Egyptian Arabic for the same task. Next, we explain how we built and annotated our resources for Gulf Arabic. Then, we describe the conducted experiments and discuss the results. Finally, we describe some future directions.

2. Related Work

The majority of the work done, so far, for Arabic sentiment analysis is MSA-based. This can be explained by the availability of some important linguistic resources, even though these are mostly not completely freely available for the open public. These resources include Part

of Speech (PoS) tagger such as MADAMIRA (Pasha et al., 2014) where some works are based on this tool either to classify opinions using its database resource (Korayem et al., 2011) or to do some rule-based sentiment analysis (Hossam S. et al., 2015). Recently, quite a large sentiment lexicon for MSA called SLSA: Sentiment Lexicon for Standard Arabic has been made freely available for research (Ramy Eskander & Owen Rambow, 2015). SLSA is compiled by using MADAMIRA MSA database to classify its entries and linking it to the SentiWordNet¹ using the English gloss to compute the polarity score (positive, negative or subjective) for each Arabic entry. However, the Arabic used in different online platforms to express opinions is frequently not MSA (J. Owens, 2013). There are many other variations and in some cases the use of the Arabic script is the only common point, taken account of the vocabulary and syntactic considerable differences between MSA and its variations. This creates a big gap between real Arabic data (web data) and the designed tools for Arabic NLP. Recently, there are some attempts to do sentiment analysis for some widely used Arabic dialects such as Egyptian (Hossam; Sherif & Mervat, 2015) and Levantine. The major difficulty, as mentioned before, is the absence of linguistic resources for different Arabic dialects which should be processed separately from each other in addition to the complexity and ambiguity of these dialects themselves.

¹ A lexical resource for opinion mining containing a list of English terms with an attributed polarity (positivity, negativity or objectivity) score, <http://sentiwordnet.isti.cnr.it>

3. Arabic Dialectal Variations

It is commonly believed by non-Arabic speakers that there is one Arabic used in the Arab world. This assumption is misleading because there are many variations of Arabic used differently depending on the region and the local culture. Actually, it is true that there is a common language that is understood by the majority of educated Arab people because it is taught at schools and it is the main language of the media in most Arab countries. This is what is known as Standard Modern Arabic (MSA) which may be seen as a simplified version of Classical Arabic² (CA). However, when it comes to daily life, MSA is rarely, if at all, used as many people find it ridiculous to use MSA with their friends or families, instead they use their own dialects. Yasir Suleiman (2013) has explained that the long rich history of the Arabic culture and the wideness of the Arab world have created a very rich linguistic variation in Arabic itself.

Most Arabic Natural Language Processing nowadays is MSA-based. However, with the rise of different social media and new technologies where people use their own dialects (informal languages) to express themselves, it is necessary to process different dialects to understand what is going on different online platforms and build applications/tools which can handle informal languages to communicate efficiently with the target users. Habash (2010) has suggested to group Arabic dialects in five (5) main groups: Egyptian, Gulf, Iraqi, Levantine and Maghrebi. This division, based on the vocabulary variation and sentential structures used in each dialect, is intended to make it easy to build resources for these dialects and automatically process them. The challenge is that these Arabic dialects are transcriptions of the spoken dialects, meaning that they do not adhere to the MSA grammar and do not have standardized spellings.

4. Gulf Arabic

Gulf Arabic is considered to be the closest dialect to the Modern Standard Arabic (MSA) for historical reasons (Versteegh, 2011). In this paper, we do not consider morpho-syntactic information because we could not find a tool which would allow us to do so. Instead, we only take account of the vocabulary, semantics and syntactic structure where a corpus study shows that the Gulf Arabic is considerably different from MSA.

4.1. Vocabulary

There are two main common linguistic phenomena in Gulf Arabic, namely the use of arabized English (English

written in Arabic script) and the use of special vocabulary. Let's take the following examples:

Example 01:

هذا المطعم انما الان- يااa

[This restaurant is an addiction. The taste of its **chicken** shawarma is very delicious and **spicy**. Its sandwiches are fable, especially they put lots of **fresh cheese**. Truly, **bravo** and these are ten stars (symbols) for you]

In this sentence, there are five (5) arabized English words (in bold in the Arabic sentence and their corresponding English translation).

Example 02:

المطعم حلو مره وكتييير اروح له ابي اروح له **الحين** بس اخر زيارة له ارتفع ضغطي ما ادري **عشان** في زحمة أو **عشان** الناس بإجازة و**متشفحين** على لأكل والمطاعم كان زحمة و**خرشة** مو مثل أي يوم

[The restaurant is **very** nice and I visit it frequently. **Now** I **want** to go there but last visit I got high blood pressure, I'm not sure if it was **because** of the long queuing or because people are in vacation so most restaurants are **full** not as any other days.]

In this example, there are seven (7) dialectal words (in bold in the Arabic sentence). Some of them are typically found in Gulf Arabic in these meanings, namely **مره**, **ابي**, **الحين** **مره**, **ابي**, **متشفحين** and others are used also in Levantine Arabic or even Egyptian Arabic such as: **بس**, **عشان**.

4.2. Semantics

There is a significant difference between word meaning in MSA and Gulf Arabic where the same word form means different thing than the known meaning in MSA which is taken always as a reference. The bolded words in the second example above can be classified into two categories:

1. Words which can be found in an MSA dictionary: this means that these words do exist in MSA. However, many of them have a conflicting part of speech (PoS) or totally a different meaning, i.e. the only similarity is the word form. For instance, the two words **الحين**, **ابي** can be found in any Arabic dictionary, **ابي** [father] as common noun and **الحين** [period of time/time] as well. This is not the case for their meanings in Gulf Arabic where **ابي** is used to mean [I want] which is a verb and **الحين** is used to mean [right now] which is an adverb.

There is another important case which we call conflicting vocabulary where the same word form has exactly the same PoS between MSA and Gulf Arabic but with the opposite meaning. This phenomena does not have any explanation apart from cultural effect since this does not happen in all Arabic dialects. Consider the following two examples:

الشاورما حقتة رهيبه وطعمها خطيبيبير.. اذا ما ذقتها فاتك الكثير انصحكم تروحونله . (1)

2 Also known as Quranic Arabic or occasionally Mudari Arabic, it is the form of the Arabic language used in literary texts from Umayyad and Abbasid times (7th to 9th centuries). It is based on the medieval dialects of Arab tribes. Modern Standard Arabic (MSA) is its direct descendant used today throughout the Arab world in writing and in formal speaking. While the lexis and stylistics of Modern Standard Arabic are different from Classical Arabic, the morphology and syntax have remained basically unchanged though MSA uses a subset of the syntactic structures available in CA.

[The shawarma is terrible and its taste is dangerous.. if you did not taste it, you missed a lot. I recommend you try the restaurant.]

بشكل عام كل شيء كويس, البطاطس رهيب وصوص الثوم اهرب والخبز الصاجي مقرمش ويجزن . (2)

[In general everything is good, the potato is terrible and the garlic sauce is more terrible and the saj bread is crispy and drives crazy]

الله يعطيكم العافية، هالمرة الاكل دافي . (3)

[God bless you, this time the food is warm.]

When normalized, the two words: خطير and رهيب [terrible and dangerous respectively] in example 1 and 2 are negative in MSA but they are used to express very positive sentiment /opinion in Gulf Arabic. The strong positive polarity is represented by making the words longer (by doubling some characters).

In example 3, the word العافية [wellness] is used in Gulf Arabic in the same meaning as in MSA which is positive. However, in other dialects, namely in North African, العافية means fire which is negative.

This shows that words in Arabic dialects are heavily influenced by the cultural environment and the context /domain they are used in.

2. Words which are not found in an MSA dictionary: these words simply are typical to Gulf Arabic meaning that they get their meanings and forms from the culture in that area, such as: خرشة, منشفين, مره from the previous example.

4.3. Syntactic structure

As shown in the above section, the PoS of words can differ between MSA and Gulf Arabic for the same word form. The difference is found also in the sentence structure. In general, Arabic allows free word order depending on the discourse context, for instance topicalization to mark emphasis. In dialectal Arabic, this free word order is even loose since there is no standard structure. The absence of punctuation makes it even harder to get the intended meanings.

There is also an important difference which can not be ignored, especially when it comes to sentiment analysis as it is crucial for getting the right polarity of an opinion. This concerns the way the negation is expressed. Actually, this needs an entire chapter given the big difference compared to MSA, however we will just include what we see important for our task. In MSA, negation is expressed in two ways:

a - Use of a negative word (the negative polarity is inherent):

Example: أكره ذلك المطعم [I hate that restaurant]

b - Use of negation particles: originally the word or the expression is positive and to reverse its polarity we use some particles depending on the PoS of the word in question. For instance, if we want to reverse the polarity of a verb, we need to take account of its tense. For instance:

لا أحب ذلك المطعم [I do not like that restaurant]

لم أحب ذلك المطعم [I did not like that restaurant]

لن أحب ذلك المطعم [I will not like that restaurant]

In this example, the main verb is أحب [like] and the negation particle depends on the tense of the verb لا for the present, لم for the past and لن for the future.

In Gulf Arabic, negation is expressed differently. In addition to the particles used in MSA, there are special particles and expressions, namely مش, مو, مهوب. The hardest case to disambiguate is when the same particle exists in both MSA and Gulf Arabic such as the particle ما which can be a relative pronoun, a question word, a negation or has even a special use. Most of these particles are separate tokens in MSA, but in dialectal Arabic, in general, people tend to write them attached to the following word.

Beside all of this, there is a huge difference when it comes to the notion of a sentence between MSA and dialectal Arabic. In Gulf Arabic, there is no a notion of sentence and no punctuation. So, a sentence can be one word, many words, a line, an entire bloc of text, etc.

5. Linguistic Resource Building

For our purpose, we collected data and compiled a sentiment lexicon containing sentiment bearing tokens. The following is a description of these resource building process.

5.1. Dataset

The biggest obstacle we faced is the lack of freely available annotated data for dialectal Arabic. To overcome this problem, we compiled our own dataset from scratch. To do so, we manually collected data from restaurant reviews specialized websites where we counted each comment or review as one document. The current corpus contains 4072 documents, without any segmentation, divided into 2647 positive, 1296 negative, 101 mixed and 28 neutral documents. The selection of this data is based on the decision that each document has to contain only one sentiment or opinion for both positive and negative categories. For instance, if a document contains more than one sentiment of the same polarity (positive/negative), the document, whatever its length, is counted as one (positive/negative) document. However, if it contains any number of different opinions given that there is at least one positive and one negative, the document is classified as mixed. Basically the mixed category can be extended as much as wanted by combining at least one document from positive category with at least one from the negative one, nonetheless this is not what we opted for because we believe that we can catch the mixed category using some rules as long as we have enough clear positive/negative categories. Concerning the neutral category, we selected documents which do not have any clear sentiment, only include use words such as purchase, try, eat, etc. We got a list of these use words or objective words from the SLSA:

Sentiment Lexicon for Standard Arabic³ where we selected words with high objective score and kept only those we thought were useful for our purpose then translated them into their equivalent in Gulf Arabic.

5.2. Feature Lexicon

In any machine learning approach, the primary task is to select representative features. To do so, we build our own feature lexicon for Gulf Arabic. Because of the absence of any reference, we started by manually collecting sentiment bearing words from the collected corpus based on our knowledge. This includes verbs, adjectives, nouns and some expressions which have an effect on the sentiment polarity. For instance, users usually do not use space between negation particle and the following verbs, adjectives or nouns. So we considered such cases as one expression and tagged it with the appropriate polarity. We also took account of the frequency of these words. We ended up with a Gulf Arabic sentiment lexicon containing 1198 positive and 894 negative entries.

Unfortunately, we could not include any morpho-syntactic features which could help a lot in the classification task. This is because of non-existence of PoS tagger for Gulf Arabic. Actually, we could have used some existing Arabic morpho-syntactic analyzer and disambiguator such as MADAMIRA⁴, but we chose not to do so because according to our experience this tool is trained on MSA and Egyptian only.

Taking Modern Standard Arabic as a reference, there is another serious issue which is the spelling inconsistency. People tend not to use space between short words or between words and particles as in this example:

(1) مچ نووون عالمطعم ورايحل هالمسا والله جربوه وقولولي رايكم فيه
[I'm mad **about** this restaurant and I'm going **there this** evening. Try it and let me know of your impression]

Considering the word boundaries, in MSA the above example should be spelt as follows (without taking account of the misspellings):

(2) مچ نووونع الماطعمو رايحله ه المساو الله جربوه وقولولي رايكم فيه

In example (1), there are 9 tokens (based on white space), but the correct tokenization is shown in the example (2) where there are 16 separate tokens. It is worth to clarify that we are not dealing here with tokenization based on the Arabic proclitics or enclitics but only with the missing white space which causes serious re/tokenization problems.

5.2.1. Feature Selection

Based on the points discussed above, we give up on using any analyzer or an MSA dictionary. Instead, we

³ A large sentiment analysis for MSA based on Aramorph and SentiWordNet. For more details see:

<http://aclweb.org/anthology/D/D15/D15-1304.pdf>.

⁴ For more information, see:

http://innovation.columbia.edu/technologies/cu14012_arabic-language-disambiguation-for-natural-language-processing-applications.

investigate the collected data and try to find the most salient features. The features we consider can be classified into three (3) main groups:

a. Sentiment bearing words: includes the Gulf Arabic sentiment lexicon entries classified into positive and negative words. We also include some orthographic information, namely words made longer to emphasize the sentiment. For example: بااااااااايخ which looks like [baaaaaaaaaad] or حلووووو مرررره [verrrry deliciouuuuus]. Such words are given the features 'strong positive word' and 'strong negative word' and attributed a high score according to their polarity and their frequency. Otherwise, words are given either 'positive word' or 'negative word' feature based on their polarity.

b. Numerical rating expressions: our corpus is retrieved from restaurant reviews websites where people use some rating expressions. These can be grouped into three groups:

-**Numerical expressions** using both Arabic and Hindi numbers such as: 100/70, 10/5, 80% 6\10 or ١٠\٧

-**Numeral expression spellings:** ١٠ من ٦ [6 out of 10]

-**Use of stars** (spelling and symbols) for instance:

هذي عشرة ★★★★★

We consider numbers above 5 if the rating is out of 10 or above 50 if the rating is out of 100 to be positive rating. We group these features into two groups: 'numerical rating positive' and 'numerical rating negative'.

3. Emoticons: people use usually lots of emoticons in social media. We collected all the emoticons in the corpus and considered only those with clear sentiment and classified them into two categories: 'positive emoticon' and 'negative emoticon'.

4. List of negation words: we find it hard to catch all the negation structures in the corpus by rules because of their inconsistency. Therefore, we collect a list of negation words tagged as 'negative particle' and give them a high negative score and very low (or zero) positive score.

6. Methods

As mentioned in the section before, Arabic dialects are variations of spoken Classical Arabic which have been influenced by different cultural factors along history and ended up in today's forms. This means that there is no standard spelling for a given dialect. Basically, this means that every spelling is correct since there is no standard spelling to compare to and incorrect taken MSA spellings as a reference. This makes it hard, if not impossible at all, to automatically process dialectal Arabic using the traditional Natural Language Processing (NLP) methods for sentiment analysis, namely grammar based approaches. That is the main reason we opt for machine learning in this paper where we use our implementation of the Naive Bayes classifier to make it flexible in order to be able to include new features and change their polarity scores as wanted. Otherwise, we would need to create rules for every user even more rules for the same user as

s/he is not always consistent, i.e. can find different spellings of the same word in the same document. Also, it is worth mentioning that we do not apply any preprocessing steps, namely normalization, spelling correction simply because there is no reference spellings. More important, we want to avoid conflicting mismatches and make use of the sentiment bearing made longer words as a good indicator of polarity intensification. Also, the tokenization we use is based on whitespace and not on any morpho-syntactic information. We remove all punctuations and stop words which are, in this case, Named Entities: names of restaurants, names of people, names of products/dishes, locations and days.

7. Experiments and Evaluation

The collected dataset contains mixed, negative, neutral and positive documents. In this section, however, we consider only classification of positive and negative documents because, at our sense, they are more challenging compared to the other categories. We use a balanced dataset (1295 positive and 1295 negative documents) divided into 80% for training and 20% for testing. We implemented a Naïve Bayes classifier (NB) based on word frequency. We conducted four (4) experiments. The purpose of these experiments is to see how well an MSA sentiment lexicon will do in Gulf Arabic sentiment classification.

7.1. Experiment A

Train the NB classifier on the training dataset and test it on the test dataset.

7.2. Experiment B

Use an MSA sentiment lexicon compiled by extracting only clearly positive and negative entries from the SLSA lexicon based on their polarity scores and tagged them as 'positive word' / 'negative word' if their corresponding polarity score ranges between 0.5-0.7 and 'strong positive word' / 'strong negative word' if their polarity score is greater than (or equals to) 0.7. We modify the NB classifier to take these features into account. For any document, all words with 'strong positive word' / 'strong negative word' feature are attributed a high corresponding polarity score (or 1) and very low (or 0) corresponding reverse polarity score. For words with 'positive word' / 'negative word' feature, we use the NB usual score. We introduce a simple rule to deal with negation such that for each clearly positive/negative token (word, emoticon or numeral) preceded by a negative particle, we reverse the polarity score.

7.3. Experiment C

Use the modified NB classifier, used in experiment B, and use the compiled Gulf Arabic sentiment lexicon instead. For any document, all words with any of these features: 'strong positive word', 'numerical rating positive'

or 'positive emoticon' are attributed a high positive score (or 1) and a very low negative score (or 0). Likewise, for words with 'strong negative word', 'numerical rating negative' or 'negative emoticon', we attribute a high negative score (or 1) and a very low positive score (or 0). Otherwise, we use the usual NB scores. We deal with negation the same way as in experiment B.

7.4. Experiment D

Use both the MSA and the Gulf Arabic sentiment lexicons as in experiment B and C.

The results of these experiments are in Table 1.

<i>Experiment</i>	<i>Accuracy (%)</i>
A	82.81
B	76.83
C	90.54
D	84.94

Table 1: Results.

8. Result Discussion

The results in table 1 clearly show the uselessness of using sentiment lexicon designed for Modern Standard Arabic (MSA) to classify sentiments in Gulf Arabic. This can be interpreted by the dropping of the classifier accuracy (76.83%) compared to the baseline (82.81%). Also, mixing both MSA and Gulf Arabic lexicons caused the classifier performance to drop (84.94%) compared to using only the Gulf Arabic sentiment lexicon (90.54%). It is important to mention that this negative effect is not caused by any misclassification of the MSA lexicon (SLSA) itself but by the fact that words are used differently in MSA and Gulf Arabic, i.e. the same word form with totally different meaning. This reinforces again our choice to treat MSA and Gulf Arabic separately and use orthographic features, namely spellings, emoticons, symbols, etc. There is also many mismatch or no-match at all between the SLSA entries and the corpus words. Of course, there is a better alternative to the non-match, namely apply a preprocessing step (normalization and spelling correction) to the corpus. However, this is not very helpful in our case because of the different usage of words between MSA and Gulf Arabic and more important diacritics are ignored. It would be also better to introduce some better rules to deal with negation. Another remarkable thing is that the use of the Gulf Arabic sentiment lexicon only has outperformed the classifier's baseline, 90.54 % compared to 82.81%.

9. Conclusion and Future Directions

Processing different Arabic dialects is crucial for understanding what is going on social media and most online platforms. Therefore, it is useless to develop systems which are only MSA-based, unless the unique

purpose is to process news, some blogs or forums content. Unfortunately, it is the case of the most current Arabic Natural Language Processing tools. To bridge this gap, we build linguistic resources, namely annotated data and sentiment lexicon for Gulf Arabic for sentiment analysis task. In this paper, we showed the importance of processing Gulf Arabic separately from MSA where the experiment showed that using an MSA sentiment lexicon had negative effects on the overall performance of the sentiment classifier of Gulf Arabic, i.e. mixing linguistic resources is not recommended. Due to the considerable differences between MSA and dialectal Arabic, we believe that it would be better to process each dialect as a stand-alone language, with its own resources, at least for sentiment analysis and opinion mining tasks in order to avoid conflicting word usage. Hence, we are planning to do the same for other dialects.

10. References

- Arfath Pasha, Mohamed Al-Badrashiny, Mona Diab, Ahmed El Kholy, Ramy Eskander, Nizar Habash, Manoj Pooleery, Owen Rambow and Ryan M. Roth. (2014). MADAMIRA: A Fast, Comprehensive Tool for Morphological Analysis and Disambiguation of Arabic. In Proceedings of the Language Resources and Evaluation Conference (LREC), Reykjavik, Iceland.
- Diab, M. (2009). Second Generation Tools (AMIRA 2.0): Fast and Robust Tokenization, POS tagging, and Base Phrase Chunking. MEDAR 2nd International Conference on Arabic Language Resources and Tools, Cairo, Egypt.
- Ditters, E. (2013). Issues in Arabic Computational Linguistics. In The Oxford Handbook of Arabic Linguistics, Oxford University Press, pp. 213—236.
- Hossam S. Ibrahim, Sherif M. Abdou and Mervat Gheith. (2015). Sentiment Analysis for Modern Standard Arabic and Colloquial. In International Journal on Natural Language Computing (IJNLC) Vol. 4, No.2.
- Mohammed Korayem, Mona T. Diab and Muhammad Abdul-Mageed. (2011). Subjectivity and Sentiment Analysis of Modern Standard Arabic. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: shortpapers, pp. 587—591, Portland, Oregon.
- Mona Diab, Nizar Habash. (2014). Natural Language Processing for Arabic and its Dialects. In EMNLP, Doha, Qatar Tutorial.
- Nizar Habash. (2010). Introduction to Arabic Natural Language Processing. Synthesis Lectures on Human Language Technologies, 3(1), 1—187.
- Nizar Y. Habash. (2010). Introduction to Arabic Natural Language Processing (Synthesis Lectures on Human Language Technologies), Edited by Graeme Hirst.
- Owens, J. (2013). A House of Sound Structure, of Marvelous Form and Proportion: An Introduction. In The Oxford Handbook of Arabic Linguistics, Oxford University Press, pp. 1—21.
- Ramy Eskander, Owen Rambow. (2015). SLSA: A Sentiment Lexicon for Standard Arabic. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP2015), <http://aclweb.org/anthology/D/D15/D15-1304.pdf>.
- Versteegh Kees. (2011). Les dialectes arabes. *Dictionnaire des langues*, pp. 336—346. Paris: Presses Universitaires de France.
- Yasir Suleiman. (2013). Arabic Folk Linguistics: Between Mother Tongue and Native Language. In The Oxford Handbook of Arabic Linguistics, Oxford University Press, pp. 264—279.