Investigating the Rate of Agreement and Disagreement of Tense and Aspect of Quranic verbs in Arabic to English Translations: Experimental Results and Analysis

Jawharah Alasmari[1,] Janet C.E. Watson [2,] Eric Atwell[3,]

[1,] School of Language, Cultures and Societies, University of Leeds, UK
& Princess Nourah bint Abdulrahman University. Saudi Arabia
[2] School of Language, Cultures and Societies, University of Leeds, UK
[3] School of Computing, University of Leeds, UK
[1] ml14jsna@leeds.ac.uk, [2] j.c.e.watson@leeds.ac.uk [3] e.s.atwell@leeds.ac.uk

Abstract

The practice and denotation of tense and aspect differ in Arabic and English, so there is a challenge when translating between the two languages, particularly when the appropriate translation depends on a range of linguistic contexts[1], comprising also the context of use. In this paper, the Qur'anic Arabic corpus of verbs is used in Arabic with their English translations by building a sub-corpus of verbs. The study uses a statistical method incorporating SPSS and Kappa feature of SPSS to investigate the rate of agreement and disagreement of Quran Verb Tense and Aspect in Arabic to English translations. The aim is to provide information that can be used to address some of the challenges that arise when translating between Arabic and English. The SPSS results indicate the highest percentage for past, present and future tenses of Quranic Arabic verbs; the progressive and perfective aspect has the lowest percentage. Kappa must is used to estimate the disagreement between translations with a stronger measure than the SPSS percent agreement calculation, while κ also takes into consideration the possibility of the agreement occurring by chance. The results show a clear disagreement between the original text, and its translations, while the agreement varies between strong and weak. This indicates that there are difficulties when translating Arabic verbs into English.

Keywords: SPSS tool, Kappa rate, Arabic verbs, English translation

1. Introduction

Arabic is part of the Semitic language family (which includes, for example, Hebrew, Aramaic and Amharic), which has a morphological system that differs from that of English or other Indo-European languages (Sawalha, 2011). To explain Arabic morphology, verb conjugations must be understood.

The verbs are conjugated in different tenses to reflect gender, plurality, voice, and other aspects. The inflectional verbal morphology of Arabic distinguishes between a suffix conjugation and a prefix conjugation. These are typically referred to by contemporary modern linguists typically as the perfect and the imperfect (Neme and Laporte, 2013). The suffix conjugation typically refers to the past tense, while the prefix conjugation typically refers to the present tense.

The grammatical categories relevant for verbs are person (first, second, third), number (singular, dual, plural) and gender (masculine, feminine). These categories are realised in the

[1] Linguistic context = discourse that surrounds a language unit and helps to determine its interpretation

suffixes of the suffix conjugation and the prefixes (and suffixes) of the prefix conjugation, as shown in the tables below:

Table 1: The Suffix Conjugation Takes the Following Suffixes

kataba "to write"			
'I write'	***a**ktub**u***	'we write'	***na**ktub**u***
'you (m.) write'	***ta**ktub**u***	'he writes'	***ya**ktub**u***
'you (f.) write'	***ta**ktub**iina***	'she writes'	***ta**ktub**u***
'you (m. pl.) write'	***ta**ktub**uuna***	'they (pl.)write'	***ya**ktub**uuna***
'you (f. pl.) write'	***ta**ktub**na***	'they (f.pl.) write'	***ya**ktub**na***
'you (dual) write'	***ta**ktub**āni***	'they (dual) write'	***ya**ktub**āni***

Table 2: The Prefix Conjugation Takes the Following Prefixes and Suffixes

kataba "to write"			
'I write'	***a**ktub**u***	'we write'	***na**ktub**u***
'you (m.) write'	***ta**ktub**u***	'he writes'	***ya**ktub**u***
'you (f.) write'	***ta**ktub**iina***	'she writes'	***ta**ktub**u***
'you (m. pl.) write'	***ta**ktub**uuna***	'they (pl.)write'	***ya**ktub**uuna***
'you (f. pl.) write'	***ta**ktub**na***	'they (f.pl.) write'	***ya**ktub**na***
'you (dual) write'	***ta**ktub**āni***	'they (dual) write'	***ya**ktub**āni***

On the other hand, In the English, there are many as sixteen different tense structures; there are tense forms such as present, past, future, and each tense has four aspectual references: simple, progressive, perfect and perfect progressive (Gadalla, 2002).

Compared to other languages such as French or Spanish, Arabic has a more complex system of morphosyntactic agreement (Holes, 2004; Habash, 2010). Alkuhlani and Habash (2001) explain this complexity in Arabic as being "partly because of its richness, and partly because of its complex morphosyntactic agreement rules which depend on features not necessarily expressed in word forms, such as lexical rationality and functional gender and number''

Employing SPSS statistical software and its kappa function, used to measure the agreement between two translations, a statistical method can be used to establish the average of agreement and disagreement between the seven translations of Quranic Arabic verbs. In order to do this, a dataset of the contexts of the Arabic verbs appearing in the verses and their seven English translations will be used.

The objective of this method is to compare and consider the highest and lowest agreement rates between the original text of the Quranic Arabic corpus, and the seven English translations by means of quantifying verb tense and aspect. The statistical evaluation of agreements helps to produce some descriptive details of Arabic verb tens and aspect, and its translations in clear and accurate ways. Using a statistical method here could thus assistance to provide a general picture of the use of verbs in the Quranic Arabic corpus, and their translations. The analysis of these details can help to improve the translation of Arabic verbs into English.

2. Literature Review

In their study, Al-Sohbani and Muthanna (2013) find that the main problem in Arabic–English translation is poor language knowledge, and inadequate and undisciplined grammar practice. To translate from Arabic into English, the correct aspectual references of each form can be delivered by considering the context, and by recognising particles controlled in the contexts.

Eades and Watson (2013) refute the idea of Arabic as possessing only two true tenses, and argue that the role of aspect is also present. They suggest that understanding the role and function of aspect is not simple. They also consider how, in some cases, the context can alter the usual syntactic effect. Verb forms can be used in different meanings or as indications that can only be understood by referring to current evidence and context.

Gadalla (2006, p.8) argues that "a good translator must fully understand the context of an Arabic tense form before attempting to render it into English. Understanding the context helps him to understand the meaning of each form, which is very important for translation". In this respect, Ibn-ʿali (2009) also states that an understanding of syntactical and morphological instruction must go hand in hand. According to Schulz et al (2000, p. 286) "[A]s the participles per se do not express a tense in Arabic language, it must be decided by the context which temporal reference is given in particular cases".

Al-Ṭabṭabā'i (1983) has identified another basis for this time-tense relationship. Al-Ṭabṭabā'i considers Arabic tense with the existence 'real tense' or nonexistence 'unreal tense' regarding the action of verb indication.

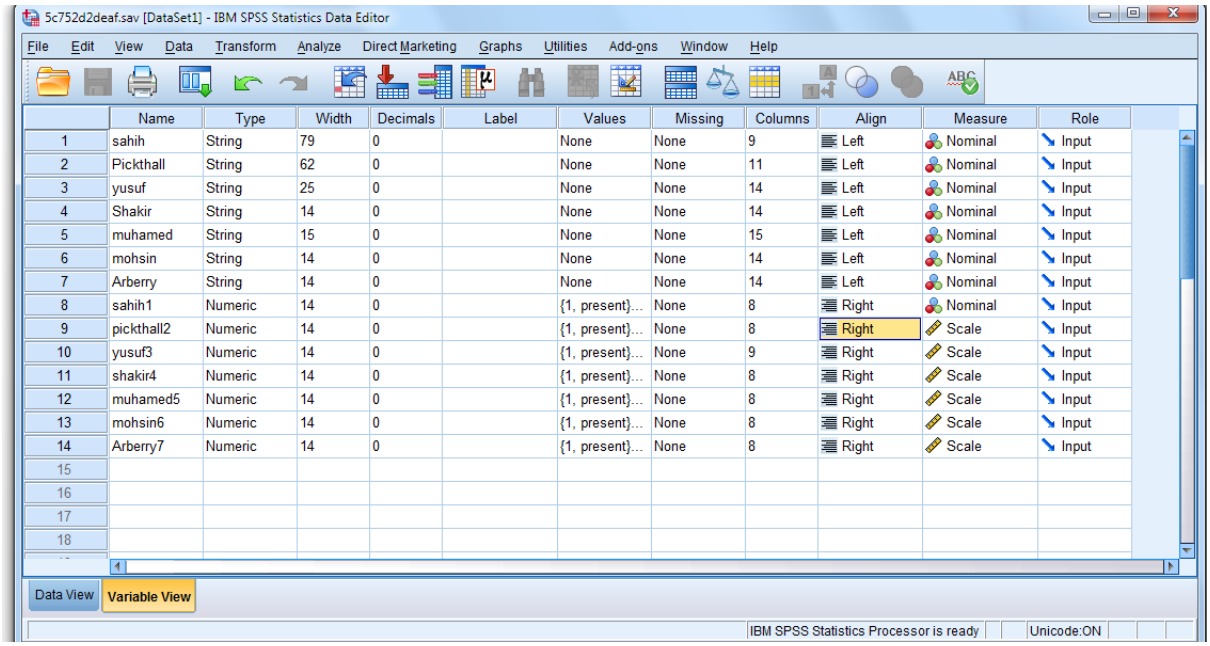
3. Methodology

In this experiment, a dataset of Quranic Arabic verbs with their seven translations were used. 1616 examples were investigated of the perfect and imperfect aspect verbs translations of the triliteral root *qāf wāw lām* ق و ل 'to say' *qāla* using SPSS software. Subsequently, Kappa feature was used to investigate the rate of agreements and disagreement of the seven translations.

A corpus syntactical and morphological analysis of the Arabic Quranic texts were used to consider statistical calculations of features such as verb aspect and mood in order to highlight

the extent of the difference between the uses of verb according to inflectional and syntactical morphology. The value was recognised depending on the Quranic corpus morphology/syntax analysis of the verbs as perfect, imperfect, imperative, subjunctive mood, jussive mood, indicative mood, and passive perfect/imperfect which are needed to consider the use of the verbs in the seven translations.

The Arabic perfect tense was translated into English as “say,” “said,” “saying,” “were saying,” “are saying,” “says,” “will say,” “shall say,” etc., which are also used to evaluate the verb translations in the software. V1, V2, V3, etc were used to refer to the seven translations. The label was divided into chapters, verb tenses, verses, and their translations. Consider Figures 1, 2, 3:



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	sahih	String	79	0		None	None	9	Left	Nominal	Input
2	Pickthall	String	62	0		None	None	11	Left	Nominal	Input
3	yusuf	String	25	0		None	None	14	Left	Nominal	Input
4	Shakir	String	14	0		None	None	14	Left	Nominal	Input
5	muhamed	String	15	0		None	None	15	Left	Nominal	Input
6	mohsin	String	14	0		None	None	14	Left	Nominal	Input
7	Arberry	String	14	0		None	None	14	Left	Nominal	Input
8	sahih1	Numeric	14	0		{1, present}...	None	8	Right	Nominal	Input
9	pickthall2	Numeric	14	0		{1, present}...	None	8	Right	Scale	Input
10	yusuf3	Numeric	14	0		{1, present}...	None	9	Right	Scale	Input
11	shakir4	Numeric	14	0		{1, present}...	None	8	Right	Scale	Input
12	muhamed5	Numeric	14	0		{1, present}...	None	8	Right	Scale	Input
13	mohsin6	Numeric	14	0		{1, present}...	None	8	Right	Scale	Input
14	Arberry7	Numeric	14	0		{1, present}...	None	8	Right	Scale	Input
15											
16											
17											
18											

Figure 1: Entering and Saving Data in the Data Editor Window

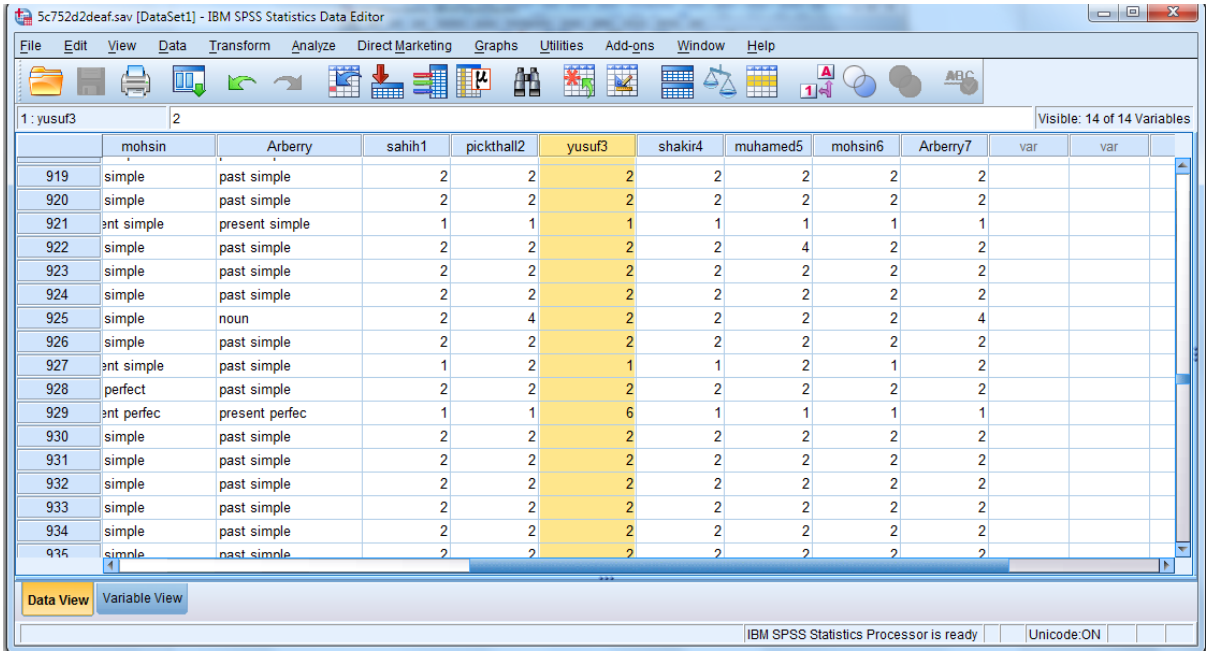


Figure 2: Entering Data in the Data View

In order to estimate the average percentages of verb translations, the descriptive statistics feature was used.

4. Results and Analysis of Using SPSS

In the statistical analysis for descriptive purposes, I compiled the following tables of examples to highlight the frequency and percentage of the verb use in the Quarnic text and the translation agreements.

Table 3: Frequencies and Percentages for the Use of the Verb *qāla* Indicating Tense, Aspect, and Mood Categories in the Quranic Text

The verb aspects/ moods			Frequency	Percent	Valid Percent	Cumulative Percent
Valid	perfect		949	58.7	58.7	58.7
	Imperfect/ indicative mood		207	12.8	9.2	67.9
	Subjunctive mood		39	2.4	2.4	92.0
	Jussive mood		20	1.2	1.2	93.2
	Passive perfect		49	3.0	3.0	96.2
	imperative		349	21.6	21.6	89.5
	Passive imperfect		3	.2	.2	96.4
	Total	1616	100.0	100.0		

Overall, the above table shows that only two aspectual forms, the perfect aspect and the imperfect aspect verb, are used to describe past actions, and actions of the present or the future. Furthermore, the indicative, jussive, and subjective moods, are indicated by only one aspectual form.

The three tenses (relating to time/ when the action happens?), the past, present and future الماضي وَالْمُضارع والمستقبل in the two aspects, are established by the Arabic verb. The progressive such as present continuous, and perfective aspects within 'event time' that gives information about a prior action are provided as well by the two forms based on the context. The impact of context effects is considered to be part of Arabic verb tenses.

The verb inflectional morphology of the perfect verb aspect forms, does not only indicate tense or aspect as the corpus revels. It is clear that there are other factors in the contexts that could affect verb tense and aspect. For examples, the past habitual aspect could be indicated by the verb structure, and the particles or auxiliary verbs. Table 4 shows the percentage of the perfect aspect verb forms translations to indicate tense and aspect. Consider the following table:

Table 4: Percentages of The translations of the Quranic Perfect Aspect Verb

Column1	**Sahih (%)**	**Pickthall (%)**	**Yusuf (%)**	**Shaki (%)**	**Muhammad (%)**	**Mohsin (%)**	**Arrberry (%)**	**Total**
future	9.3	3.6	8.0	8.7	7.7	8.6	7.1	7.57
future passive	0	0	0	0	.1	0	0	0.02
infinitive	0	0	0	0	0	0	0	0
meaning	.1	.1	.8	.1	4.3	.3	0	0.83
Noun /Participle	.2	2.1	1.3	.4	4.7	.8	2.3	1.70
past	0	0	0	0	0	0	0	0
past continuous	0	0	0	0	0	0	0	0
past passive	0	0	0	0	.2	0	0	0.03
past perfect	1.1	.9	.9	.8	1.6	.9	.8	1.02
past perfect continuous	0	0	0	0	0	0	0	0
past simple	74.4	71.3	71.3	73.4	65.6	73.8	73.8	71.96
Perfect Would- have been- used to	0	.1	.1	.1	0	0	0	0.05
present continue	0	0	0	0	0	0	0	0
present passive	.1	.1	.1	.1	.1	.1	.1	0.11
present perfect	.9	.5	.3	.2	5.6	.4	.6	1.23
present simple	13.9	21.2	17.1	16.0	10.0	15.0	15.3	15.49
								100.00

The results indicate that the Arabic perfect verb aspect, which is usually built with suffixes, is used to indicate multiple tenses/aspects, as the translations suggest. In the English translations, it can be considered that as many as sixteen different structures; there are tense forms such as present, past, future, and each tense has aspectual references such as: simple, progressive, perfect and perfect progressive. The simple past is used in 71.96 percent of cases, signifying the highest percentage; and the simple present in 15.49 percent of cases. The future tense is realized in 7.57 percent of cases, and indicates the lowest percentage of using a perfect verb. Perfective/imperfective aspects are indicated by the prefix conjugated verb as past perfect with 1.02 %, perfect models with 0.05 %, and present perfect with 1.23 percent. Other verb terms such as infinitives, nouns, or participles account for only 4.7 percent as the highest percentage of using a meaning.

An investigation of the corpus translations reveals that Quranic Arabic verb structures are used to indicate both time (absolute/relative) and action. The progressive and the perfective aspects are expressed by the perfect aspect, which could also express the absolute and the relative past tenses, or may express a completed and uncompleted action. Additionally, the results show that the present perfect that is used to express a past event that has present consequences is established as well.

5. Results and Analysis of Using COHEN'S KAPPA Statistic

Cohen's kappa (also referred to as kappa) was used in the statistical analyses of agreement across the seven translations. While SPSS was used to compare between the original text of Quranic Arabic corpus, and the seven translations to show their compatibility and the extent of differences in the expression of the verb in terms of tense/aspects between them, Kappa was used to estimate the agreement and disagreement between translations with a stronger measure than the SPSS percent agreement calculation, while κ also takes into consideration the possibility of the agreement occurring by chance. Using Kappa helps to show strength or weakness of agreement between the translations, which may be further evidence indicating the differences between them. This disagreement may be the result of various challenges translators encounter when translating Quranic Arabic verb tense/aspect.

Kappa is primarily used to measure the agreement between two individual sounds of the same type. It is generally thought to be a stronger measure than a percent agreement calculation, as κ also subtracts out agreement due to chance. Accordingly, Steven explains (online), "Kappa measures the percentage of data values in the main diagonal of the table and then adjusts and these values for the amount of agreement that could be expected due to chance alone."

The kappa value is equal to 1 or less. Perfect or complete agreement is a value of 1, while zero or less than 1 indicates moderate or low agreement. Consider the following tables which provide details of the process of the experiment and its results. The tables 5, and 6 show the difference of results between kappa and the percent agreement calculation:

Table 5: The Percentage calculation of Agreement between the translations

	Sahih International	Pickthall	Yusuf Ali	Shakir	Muhammad Sarwar	Mohsin Khan	Arberry
Sahih International	100						
Pickthall	76.88442	100					
Yusuf Ali	71.8593	71.8593	100				
Shakir	79.8995	75.8794	79.39698	100			
Muhammad Sarwar	61.80905	63.31658	64.82412	67.8392	100		
Mohsin Khan	73.36683	76.88442	82.91457	79.8995	71.8593	100	
Arberry	77.38693	71.35678	70.35176	75.8794	62.31156	73.86935	100

Table 6 Calculation of Cohen's Kappa

	Sahih International	Pickthall	Yusuf Ali	Shakir	Muhammad Sarwar	Mohsin Khan	Arberry
Sahih International	1						
Pickthall	0.62	1					
Yusuf Ali	0.55	0.53	1				
Shakir	0.68	0.67	0.66	1			
Muhammad Sarwar	0.41	0.41	0.45	0.49	1		
Mohsin Khan	0.58	0.62	0.73	0.68	0.57	1	
Arberry	0.65	0.54	0.53	0.62	0.42	0.59	1

Subjects = 199
Raters = 7
Kappa = 0.566

As a data corpus of the seven translations is used to compile the results shown above, it can be noted that there is a moderate agreement between the Sahih International and other translations, and a substantial agreement with the Shakir translation, which has the highest value at 0.68. The lowest value of agreement, 0.41, is found with the Muhammad Sarwar translation, which yields moderate agreement. According to the value of agreement between the Muhammad Sarwar translation and every other translation, the results provide the lowest value, with a moderate agreement between 0.57 and 0.41.

Overall, the results show a clear disagreement between the translations, while the agreement varies between strong and weak. This indicates that there are difficulties when translating Arabic verbs into English.

6. Discussion of the finding

According to the results, it is clear that the descriptive statistical analysis has shown that while Arabic has only two aspectual forms, perfect and imperfect, English has more grammatical classes for tenses. Additionally, there is variation in the agreement between the seven translations in terms of verb tense/aspect. It was clear that the whole context and technical attentions by the translators have an effect on the Quranic verb translations.

In Arabic, the use of inflectional morphology aspects and the use of syntactical morphology may had a reasonable effect on verb tense and aspect, indicating that some translators were seemingly unable to translate verbs by recognising such as the prefixes and affixes that are added to the verb forms, and some particles or auxiliary forms.

Tucker (2010) confirms when suggesting that English tenses do not follow the same patterns as Arabic tenses. Shamaa (1978, pp. 32–33) also discusses potential reasons for the difficulty encountered when translating Arabic tenses into English: "Temporal contrasts in Arabic are less systematic, i.e., they are not clearly marked by verb forms. […] Temporal reference in Arabic is expressed by means of verb forms in conjunction with time adverbials and other lexical items. It is, however, the context which […] finally places the action or event in its true temporal and aspectual perspective. But since context may not provide the same clear-cut and easy determinations afforded by some European [e.g. English] tense systems, it is therefore a source of occasional ambiguity".

It is without doubt that each translator has his/her own way of translating, as the results show. Muhammad Sarwar translations, for example, may base meaning in most translations cases on contexts without recognizing the role of grammatical elements such as particles and auxiliary verbs which can affect the verb tense/aspect in the context. In this regard, Gadalla (2006, p. 244) makes the following observation "Certain verbs such as كان *kāna* 'to be' and certain particles such as قد *qad* 'already' combine with these two forms of the verb to convey various meanings''. Thus, one of the major problems facing English–Arabic translators is identifying the Arabic verb form and the verbs or particles that can be combined within it in order to convey a particular English tense. Therefore, morphological and syntactic issues have to be addressed when translating Arabic verbs.

7. Conclusion

The descriptive statistical analysis has shown that while Arabic has only two aspectual forms, perfect and imperfect, English has more grammatical classes for tenses. It was clear that the whole context and technical attentions by the translators have an effect on the Quranic verb translations. From the experiment discussed here, it can be concluded that using a statistical method is a useful way to collect, analyse, and make inferences from the described data. Choosing appropriate English translations of Quranic Arabic verb tense and aspect using a quantitative approach and statistical methods can supplement the qualitative analyses of the data. Further, the data can be classified and counted in order to explain the observed facts, which can then be used to improve the translation of the suffixes and the prefixes conjugation of verb systems between the two languages.

8. References

Alasmari, J., & Atwell, E., Watson, J.2017. Using the Quranic Arabic Corpus for Comparative Analysis of the Arabic and English Verb Systems. International Journal on Islamic Applications in Computer Science and Technology. Vol 5. Available at: http://www.sign-ific-ance.co.uk/index.php/IJASAT/article/view/1683

Alasmari, J., & Atwell, E., Watson, J.2017. A Comparative Analysis Of Verb Tense And Aspect In Arabic And English Using Google Translate. International Journal on Islamic Applications in Computer Science and Technology. Vol 5. Available at: http://www.sign-ific-ance.co.uk/index.php/IJASAT/article/view/1683

Alkuhlani, A. and Habash, N. A Corpus for Modeling Morpho-Syntactic Agreement in Arabic: Gender, Number and Rationality. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics. Portland, Oregon. Pages 357–362.

Al-Sāmrā'i, I. 1983. Al-Fiçl Zamānuh wa 'abnīyathu. Bagdad. Maṭbaçāt Al-çĀni.

Al-Sohbani, Y. & Muthanna, A. 2013. Challenges of Arabic-English Translation: The Need for Re-Systematic Curriculum and Methodology Reforms in Yemen. Academic Research International, 4(4), p.442.

Eades, D., & Watson, J. C. 2013. Tense and Aspect in Semitic: A Case Study Based on the Arabic of the Omani Šarqiyya and the Mehri of Dhofar. Boston: Brill.

Gadalla, A. H. 2006. Arabic imperfect verbs in translation: a corpus study of English renderings. Meta: Translators' Journal, 51(n° 1), pp. 51–71.

Gadalla, H. & Abdel-Hamid, A. (2000). Genitive constructions in English and Arabic: a contrastive study. The Bulletin of the Faculty of Arts, 6. Assiut University, Egypt. pp. 1-64.

Habash, N. Y. (2010). Introduction to Arabic Natural Language Processing. UK: Morgan & Claypool Publishers.

Holes, C. 2004. Modern Arabic: Structures, Functions, and Varieties. Washington, D.C.: Georgetown University Press.

Ibn Al-'Aṭ̃ir, Ḍ.1983. T: Aḥmad, A., and Badawi, T. Al-Maṯal Al-Sā'ir fi 'Adab Al-Kātib wa Al-Šāçir. Dār Al-Rafāçi.Al-Riyaḍ. Al-Mujalad 2.

Neme, A. A., & Laporte, E. (2013). Pattern-and-Root Inflectional Morphology: the Arabic Broken Plural. Language Sciences, Elsevier, 40, pp.221-250.

Sawalha, M. S. S. (2011). Open-source Resources and Standards for Arabic Word Structure Analysis: Fine Grained Morphological Analysis of Arabic Text Corpora. PH.D. Thesis, University of Leeds.

Schulz, E., et al. 2000. Standard Arabic. Cambridge University Press.

Shamaa, N. 1978. A Linguistic Analysis of Some Problems of Arabic to English Translation. PhD. Dissertation, Linacre College, Oxford.

Simon, S. What is a Kappa coefficient? (Cohen's Kappa). [Online]. Available at: http://www.pmean.com/definitions/kappa.htm [Accessed 2 March 2017].

Tucker, A.M. 2010. The Morphosyntax of the Arabic verb: toward a unified syntax-prosody. University of California, Santa Cruz, Qualifying Exam Paper.95064-1077.