The Effect of Vocalization on The Translation of Arabic Sentences into English

Asst. Prof. Dr. Marwah Kareem Ali

Department of Translation, College of Arts, Tikrit University
Salahuddin, Iraq

A B S T R A C T

In Arabic, vocalization, also known as tashkil and diacritization, refers to signs that represent the vowels and some other phonetic and phonemic entities, which are placed above, below, or within the letters. As stated in literature, Arabic sentences can be read and written without diacritics, but their intended syntactic and semantic structures may not be easily understood. This is because diacritics can change the form, function and meaning of sentence elements. This in turn leads to changing the type of sentence and its significance. Thus, diacritics are important as they play a great role in identifying the intended syntactic and semantic structures of Arabic sentences. However, in case of translation, the absence of vocalization or diacritization causes problems in determining the sentence form, type of pronouns used in the sentence, as well as its meaning. Therefore, this study aimed at identifying the effect of vocalization or diacritics on providing the appropriate translation for Arabic sentences. In this concern, a descriptive analytic method was used to analyze the data of study. The study sample consisted of (6) Arabic sentences of different structures (diacritic and non-diacritic) distributed to (5) subjects to translate them into English. Hence, pre- and post-tests were applied to compare the effect of vocalization on the translation. In pre-test, the sentences were not vocalized; whereas in post-test, they were vocalized. The results revealed the significant role played by vocalization on determining the intended syntactic and semantic structures and then providing the appropriate translation for these structures.

K E Y W O R D S

1.1. Introduction:

Vocalization, also known as tashkil and diacritization, refers to signs representing vowels and some other phonetic and phonemic entities. These entities are known diacritics or diacritical markers placed above, below, or within the letters. Vocalization has a significant effect on and plays an important role in stabilizing the oral transmission of some reading traditions of texts (Morag, 1972). According to Maroun (2017: viii), in Arabic orthography, diacritics "are typically omitted from written texts, thereby making many Arabic words phonologically and semantically ambiguous. Such words are known as heterophonic homographs and are associated with different pronunciations and meanings". In Arabic, letters represent only consonants and long vowels, while other phonological characteristics, primarily short vowels, are represented by optional diacritics (Hallberg, 2022).

Despite the fact that there are several studies investigating the system of vocalization/ diacritics in Arabic, these studies dealt with this system from different perspectives including reading comprehension, differentiating this system among languages, as well as designing electronic models to identify diacritics in machine translation. However, according to the researcher's knowledge, there is no previous study dealt with this system to identify its effect on the translator's awareness of its role in identifying the intended syntactic and semantic structures of Arabic sentences. Therefore, this study aims at identifying the effect of this system on translators to provide an appropriate translation. To achieve this objective, Arabic sentences with/ without diacritics are presented to subjects specialized in translation to render them. It is hoped that this study would add to the existing literature on the issue under study through dealing with it from a different perspective, which is translation. In addition, it would be of significance for academics specialized in translation, linguistics and language teaching to benefit from its content and results and enhance their existing knowledge on such topic.

1.2. Literature Review:

This section reviews the literature on subjects relevant to the issue under study. It highlights the reading and writing systems of Arabic language, vocalization / diacritization in Arabic as well as previous studies related to this issue.

1.2.1. Arabic Reading and Writing Systems:

As stated by Maroun (2017: 10), "Arabic is the fifth most common spoken language in the world and its script is the second most widely used one after Roman". It is a primarily consonantal writing system, providing vowel information through using diacritics above or below the word. Diacritics have a considerable role in clarifying the phonological features and increasing the visual complexity of words. Typically, in written texts, vowels are omitted; therefore, readers must rely on other means like context in order to understand those texts.

In Arabic, writing system has a number of diacritics (taškīl) primarily indicating short vowels. "These diacritics are used to varying extents, giving a form of orthographic variation potentially affecting every word in a text and various aspects of the reading process" (Hallberg, 2022: 1). In normal reading, the use of diacritics is typically restricted to primary school education. During the fourth to the sixth

grades in primary education, the main purpose for using diacritics is to initiate the learner on how to read without them. The vowel diacritics are not used in most of the printed works in the Arab world. However, they are added to a heterophone in case if the context does not sufficiently disambiguate it (Schultz, 2004).

Letters in Arabic language only signify consonants and long vowels. After the beginning of Islam, Arabic was evolving into a religious, administrative, and literary language. Accordingly, a system of additional signs was established to demonstrate other phonological features, primarily short vowels. This system of signs was employed to facilitate the recitation of the Glorious Quran. These signs were placed above and below the letters. Arabic writing system of letters has two categories. The first category is cursive, referring to the graphical connection of letters that changes form based on their position in a word. In fact, most letters have four forms: isolated, initial, medial, and final (e.g., ب بـ ـبـ ـب, and بـ). The second category includes optional diacritics, which indicate the phonological information missed in the first category. These diacritics are pronounced after the letter to which they are added (e.g., بَ ba, بُ bu, and بِ bi) (Hallberg, 2022).

For readers who are novice, it is difficult to read undiacritized text, as only a partial representation of the word's phonological form is provided. A text with no diacritics also has many homographs. For instance, the word (درس) (drs) may indicate these words: (دَرَسَ) darasa 'he studied' (past active); (دُرِسَ) durisa 'it was studied' (past passive); (دَرَّسَ) darrasa 'he taught' (past verb of teach); (دَرس) dars 'lesson' (noun); or (دَرِّسْ) darris 'teach (as imperative verb)'. For correctly identifying words, skilled readers depend heavily on "top-down strategies" based on context and their knowledge of "the complex Arabic morphology" (Abu-Rabia, 2007). By using diacritics, words can be disambiguated, thus simplifying their identification for novice readers (Saiegh-Haddad and Henkin-Roitfarb, 2014). Printed Arabic words are visually, orthographically and psycho-linguistically opaque and "map linguistic information that is absent from and hence not used in natural lexical processing in Spoken Arabic" (Hallberg, 2022: 4).

1.2.2. Vocalization and Diacritics in Arabic:

In Arabic, words are written mainly as consonantal roots; consonants are usually represented by letters excepting (3) letters that indicate both consonants and long vowels, /ا/-/aː/, /و/-/ uː /, /ي/-/iː/. While the three short vowels act as diacritical marks, /a/ and /u/ placed above the letter and /i/ placed below it. Vowels can be doubled at the end of nouns and adjectives that are indefinite (Maroun, 2017). Moreover, there are four other diacritics as follows:

- Sukoon denotes the absence of vowels after a consonant.
- Shaddah denotes the reduplication of letter.
- Maddah, placed above the mixed letter /ا/-/a/, denotes a glottal stop with a long vowel ā.
- Hamzah denotes a glottal stop.

In diacritized Arabic, the word's phonological information is complete as vowels are represented as diacritics. However, vowels are not represented in undiacritized Arabic. Thus, fluent readers have to derive the phonological features and the pronunciation of word through semantic context. This

is attributed to the fact that undiacritized written word can be phonologically ambiguous and could represent more than one word with a different pronunciation and meaning. Such written words are defined as heterophonic homographs. Sometimes, the same sequence of letters could denote up to eight words, which are phonologically and semantically different. Because of this homograph phenomenon, the appropriate pronunciation has to be inferred by the reader in accordance with the context. For example, the sequence علم/ ʕlm/ may be pronounced as (عِلم) /ʕilm/ 'science' or as (عَلَمٌ) /ʕalam/ 'flag' (noun) or as (عَلِمَ)/ʕalima/ 'he knew' (past tense of verb) etc. (Maroun, 2017). Additionally, diacritization is used for indicating the grammatical function of words within a sentence. In other words, the word's final letter in a written sentence is diacritized in accordance with its grammatical function, requiring the proficient reader to have advanced phonological and syntactic skills (Abu-Rabia, 2002).

There are (8) optional diacritics in Modern Arabic orthography, two of them are sukoon and shaddah, explained above, and (6) of them have two natural groups. The first group includes (فتحه) fatḥa, (ضمه) ḍamma, and (كسره) kasra, representing either these short vowels /a/, /u/, and /i/, or their long variants when used in combination with specific letters, which cannot be followed by another diacritic. The second group includes (التنوين) nunation diacritics fatḥatān, ḍammatān, and kasratān, representing the word final inflectional morphemes /an/, /un/, and /in/, taking the form of the doubled associated short vowel diacritic (Hallberg, 2022). They are only used above the final letter of word, typically signifying case inflection. Fatḥatān is also used as a derivational suffix forming adverbs (Hallberg, 2016; Ryding, 2021). Only shaddah combines with other diacritics on the same letter, but not with sukoon, as this case will indicate a sequence of three consonants, which is not acceptable in Arabic phonotactics (Hallberg, 2022).

1.2.3. Previous Studies:

This section presents previous studies in relation to vocalization / diacritization and translation. In this regard, when reviewing the literature, several studies were conducted in relation to vocalization/ diacritization. However, these studies were not about investigating the effect of vocalization on translation. Therefore, this makes this study novel in terms of investigating the effect of vocalization/ diacritization on realizing the correct syntactic and semantic structures of Arabic sentences, which contributes to providing an appropriate translation for them. The highly related studies are presented in the following paragraphs chronologically.

In (2007), Diab, Ghoneim and Habash studied Arabic diacritization in the context of statistical machine translation, aiming at investigating the impact of Arabic diacritization on statistical machine translation (SMT) through defining several diacritization schemes ranging from full to partial diacritization as well as designing models for identifying diacritics. They found that none of the partial diacritization schemes significantly varies in performance from the no-diacritization baseline despite the increase in the number of types in the data.

In 2017, Maroun presented a PhD dissertation on diacritics and the resolution of ambiguity in reading Arabic, aiming at investigating how proficient readers of Arabic process diacritics, and how they understand heterophonic homographs with and without diacritics. Hence, six experiments

were reported on this issue. Results showed that diacritics improved the comprehension of ambiguous words (i.e., heterophonic homographs) without affecting their speed, but that they had no effect on the comprehension of unambiguous words while slowing them.

Hallberg (2022) carried out a study on principles of variation in the use of diacritics (taškīl) in Arabic books. The study represented the first empirical investigation into the variation in how Arabic diacritics are used. It employed quantitative corpus linguistic methods to explore diacritization in a 72-million-word corpus consisting of book-length texts of various genres. It revealed that children's literature and poetry were found to vary considerably in the number of diacritics used, while books of normal prose fall within a narrow range of limited use of diacritics.

1.4. Research Method:

This study is based on a qualitative research method. In addition, the content analysis method is used to categorize data under analysis according to their content, thus facilitating their identification for readers. It focuses on providing a descriptive analysis for data collected represented by (6) Arabic sentences of different structures to be translated into English. The study sample consists of (5) subjects, lecturers and MA candidates, specialized in translation to translate the sentences. These subjects are numbered sequentially as (T1, T2, T3... etc.). The study utilizes the pre- and post-tests to compare the effect of vocalization on translation. In pre-test, the sentences are presented to subjects without diacritics. Whereas in post-test, the sentences are presented with diacritics. To achieve the objective of this study, the results of pre- and post-test are compared to determine the effect of vocalization on the process of translation as well as finding similarities and differences between them. As there are more than one term referring to the issue under study, thus the two terms (vocalization) and (diacritization) are used interchangeably throughout the study.

1.5. Data Analysis:

This section provides the analysis of data collected, which consisted of (6) Arabic sentences. These sentences are categorized into three subsections related to the main syntactic and sematic structures that are highly affected by vocalization. In each subsection, there is a brief introduction and then the sentences under study are presented as follows: the source text (henceforth, ST) is presented at first, and then followed by the target texts (henceforth, TTs), noting that sentences along with their translations in both tests are placed in tables in order to be easily recognized by readers. Each subsection has two parts of analysis. The first part provides the non-diacritic sentences with their translations. While the second part provides the diacritic sentences with their translations. After that, the two translations for each sentence are discussed and compared to determine the effect of vocalization on the process of translation, as well as highlighting the similarities and differences between them. As a final step after discussing the translations, a suggested translation is provided for each ST.

1.5.1. Active or Passive Voice?

This section discusses the translation of Arabic sentences with and without diacritics in relation to their syntactic and semantic structure of active / passive voice. Since determining whether

the sentence is active or passive is highly affected by the use of diacritics, this in turn affects the appropriateness of its translation.

1.5.1.1. Pre-Test Sentences:

This section presents the non-diacritic sentences along with their translations, as shown in table (1):

Table (1): Translation of non-diacritic sentences

No.	ST	TTs
1	انتخب الطالب قدوة	T1: The student elected a model student.
		T2: The student elected a model student.
		T3: The student elected a model student.
		T4: The student elected a model.
		T5: The student elected a model.
2	عين احمد مديرا للشركة	T1: Ahmed appointed a director for the company.
		T2: Ahmed appointed a manager for the company.
		T3: Ahmed appointed a company manager.
		T4: Ahmed appointed a director to the company.
		T5: Ahmed appointed a director to the company.

1.5.1.2. Post-Test Sentences:

In this section, the diacritic sentences are presented with their translations, as shown in table (2):

Table (2): Translation of diacritic sentences

No.	ST	TTs
1	انتُخِبَ الطالبُ قدوةً	T1: The student was elected as a model.
		T2: The student was elected as a model student.
		T3: The student was elected as a model student.
		T4: The student was elected as a model.
		T5: The student was elected as a model.
2	عُينَ احمدٌ مديراً للشركةِ	T1: Ahmed was appointed as the company's director.
		T2: Ahmed was appointed as a manager of the company.
		T3: Ahmed was appointed as a company manager.
		T4: Ahmed was appointed as the director of the company.
		T5: Ahmed was appointed as a director to the company.

Concerning the first ST (انتخب الطالب قدوة) in pre-test, it is clear that all translators considered it as an active voice sentence, thus they translated it as (The student elected a model) with slight differences in relation to the translation of word (قدوة), which is translated as (a model student) by T1, T2 and T3. In fact, these translations are accepted, as they look correct syntactically and semantically. Since there is no diacritic, both sentences can be translated as being active not passive as well as vice versa, so both readings can be correct. The same applies to the second ST (عين احمد مديرا للشركة), which is translated as (Ahmed appointed a director for the company), with different translations for (مديرا للشركة) as (a manager for the company), (a company manager), and (a director to the company) by T2, T3, T4 and T5, respectively.

However, the situation is totally different in post-test, as all translators agreed on the same translation after presenting the sentences with diacritics, contributing to specifying the type of sentence

voice as being passive not active. Thus, when providing diacritics, the syntactic and semantic structures become clear and accept only one correct translation. Hence, the first ST (انتُخِبَ الطالبُ قدوةً) is translated as (The student was elected as a model), while the second one (عُينَ احمدٌ مديراً للشركةِ) is translated as (Ahmed was appointed as a manager of the company). The translations in post-test are more appropriate as they represent the intended syntactic and semantic structures of ST.

The suggested translations for these STs are as follows:

1. The student was elected as a model.
2. Ahmed was appointed as a director of the company.

1.5.2. A Declarative, an Interrogative, or an Imperative Sentence?

In this section, sentences that could refer to different types of sentences at the same time are presented along with their translations and discussion. Hence, without diacritics, a sentence could be understood as being a declarative, an imperative or interrogative one. Therefore, this issue is taken into account to examine participants' awareness of it when translating the sentences.

1.5.2.1. Pre-Test Sentences:

The non-diacritic sentences along with their translations are presented in table (3).

Table (3): Translation of non-diacritic sentences

No.	ST	TTs
1	سجل أسماء الطلبة الحاضرين في الدرس	T1: He registered the attendance students' names.
		T2: A record of the names of attending students.
		T3: He recorded the names of the students who attended the class.
		T4: The names of the students attending the lesson were registered.
		T5: He recorded the names of the students who were present at the lecture.
2	حفظت القصيدة	T1: The poem was memorized.
		T2: I learned the poem by heart.
		T3: I memorized the poem.
		T4: The poem is memorized.
		T5: I memorized the poem.

1.5.2.2. Post-Test Sentences:

In this section, the diacritic sentences are presented with their translations in table (4).

Table (4): Translation of diacritic sentences

No.	ST	TTs
1	سَجِّلْ أسماءَ الطلبةِ الحاضرين في الدرسِ	T1: He registered the attendance students' names.
		T2: Write down the names of attending students.
		T3: The names of the students who attended the class were recorded.
		T4: Register the names of the students attending the lesson.
		T5: Record the names of the students who are present at the lecture.
2	حَفَظْتَ القصيدةَ؟	T1: I memorized the poem.
		T2: The poem was learned by heart.
		T3: Was the poem memorized?
		T4: Memorized the poem?
		T5: Do you memorize the poem?

Table (3) shows the translation of sentences with no diacritics. In this section, the issue differs from that considered in the previous section. Here, the two sentences seem to be declaratives when there are no diacritics. However, translations shown in table (4) for sentences with diacritics reveal the opposite. As for the first ST (سجل أسماء الطلبة الحاضرين في الدرس) in pre-test, the word (سجل) could have three possible syntactic and semantic structures due to the absence of diacritics. It could be a noun translated as (record or list), or it could verb in the past tense of a verbal sentence translated as (wrote down), or it could be a verb of an imperative sentence and then is translated as (write down). Accordingly, there are different translations for this verb according to these meanings, as all translators translated it as a verb in past tense (recorded and registered), except for T2 who translated it as an indefinite noun (a record of). The same applies to the second sentence (حفظت القصيدة). It could be a declarative sentence informing of something or a declarative sentence used for interrogating about something according to how it is vocalized. In such cases, the problem can be solved merely by diacritizing the sentence.

In post-test, the sentences are presented with diacritics; therefore, their syntactic and semantic structures become clear, affecting their translations as well. As for the first sentence (سَجِّلْ أسماءَ الطلبةِ الحاضرين في الدرسِ), most of the translators translated it as being an imperative sentence with slight differences in relation to the verb and other sentence elements. While only two translators translated it as a declarative sentence even after diacritizing it. Hence, it was translated appropriately as (Write down the names of attending students), (Register the names of the students attending the lesson) and (Record the names of the students who are present at the lecture) by T2, T4 and T5, respectively.

Unfortunately, most of the translators failed to provide an appropriate translation for the second sentence (حَفَظْتَ القصيدةَ؟) though it was presented with diacritics and a question mark to clarify its purpose of interrogation. This is because some of them did not use the correct equivalent for the intended purpose of it. In fact, this sentence is declarative in form and interrogative in function. Though it does not include any question word, the way it is vocalized through using diacritics makes it a question. Thus, it can be interpreted as (هل حَفَظْتَ القصيدةَ؟). Hence, T1 translated it as a declarative sentence (I memorized the poem), considering that the attached pronoun (ت) (Ta'a) denotes first person singular (I) as being the doer of the action. While T2 translated it as a passive voice and as an agentless sentence (The poem was learned by heart). As for T4, the translation is inappropriate syntactically and semantically as it lacks the grammatical structure that makes it understandable. It is translated as an agentless sentence with a verb in past tense and ending with a question mark (Memorized the poem?). Though T3 translated it as a question (Was the poem memorized?), but it is a question in passive voice and the ST is not in a passive voice as it is a declarative sentence directed from someone to another one asking if he memorized the poem. Concerning T5, the translator did one mistake in changing the tense of the verb from past to present (Do you memorize the poem?).

The suggested translations for these STs are as follows:

1. Write down the names of students attending the lesson.

2. Did you memorize the poem?

1.5.3. Neutral, Masculine or Feminine Reference?

In addition to the above structures, sentence reference using attached or separated pronouns can be understood as being neutral, masculine or feminine. This in turn causes problems in relation to the selection of appropriate translation. Thus, without diacritics, the intended reference may not be clear for translator.

1.5.3.1. Pre-Test Sentences:

The non-diacritic sentences along with their translations are presented in table (5) as follows:

Table (5): Translation of non-diacritic sentences

No.	ST	TTs
1	سمعته يتكلم عنك بخير	T1: I heard him mentioned your good qualities.
		T2: I heard him speak well of you.
		T3: I heard him speaking highly of you.
		T4: I heard him speak well of you.
		T5: I heard him talking about you in a good way.
2	غسلت سيارتي	T1: I washed my car.
		T2: I washed my car.
		T3: I washed my car.
		T4: I washed my car.
		T5: I washed my car.

1.5.3.2. Post-Test Sentences:

The diacritic sentences along with their translations are presented in table (6) as follows:

Table (6): Translation of diacritic sentences

No.	ST	TTs
1	سَمِعتْهُ يتكلمُ عنكَ بخيرٍ	T1: I heard him mentioned your good qualities.
		T2: You heard him speak well of you.
		T3: I heard him speaking highly of you.
		T4: I heard him speak well of you.
		T5: She heard him talking about you in a good way.
2	غَسَلَتْ سَيارَتي	T1: She washed my car.
		T2: She washed my car.
		T3: She washed my car.
		T4: She washed my car.
		T5: She washed my car.

This section discusses the importance of diacritics in identifying the deictic reference of the sentence. Through observing the translations of the first sentence (سمعته يتكلم عنك بخير) in pre-test with no diacritics, it is clear that all translators used first person singular pronoun (I) as a subject for this sentence. They translated it as follows: (I heard him mentioned your good qualities), (I heard him speak well of you), (I heard him speaking highly of you) and (I heard him talking about you in a good way), for T1, T2/T4, T3 and T5, respectively. Then, they applied the same rule to the second sentence (غسلت سيارتي), thinking that the attached pronoun (تاء) is (تاء الفاعل) and interpreting it

as being related to first person singular pronoun (I); therefore, all of them translated it as (I washed my car). In undiacritized written texts, this pronoun (I) has a neutral reference as the identity of the speaker is not known whether it is masculine or feminine because it can refer to both of them. Unlike spoken texts, even if there are no diacritics, its reference is known from the speaker himself or herself.

Hence, in a written text when mentioned alone, the reference of this pronoun is neutral, as we do not know the identity of the speaker whether it is (he) or (she). However, the use of diacritics plays a great role in identifying this identity as being either masculine or feminine. Accordingly, when presented with diacritics in post-test, some changes happen to the translation of both sentences. As for the first sentence (سَمِعَتْهُ يتكلمُ عنكَ بخيرٍ), almost all translators failed to provide the intended deictic reference of this sentence, translating it as being related to (I) and (You), except for T5 who correctly translated it as being feminine reference (She). The problem in this sentence is that the translators may not consider the base form of the verb (سَمِعَتْهُ), which is (سَمِعَ) from (فَعِلَ), and the pronoun (تاء التأنيث الساكنة) is attached to it indicating feminine reference. Moreover, there is another attached pronoun, which is (الهاء) referring to a hidden element in the sentence. This sentence is interpreted as (هي سَمِعَتْ علي يتكلمُ عنكَ بخيرٍ).

When translating the second sentence after presenting it with diacritics (غَسَلَتْ سَيارَتي), all translators succeeded in providing the appropriate reference, which is feminine, using third person singular pronoun (She), and all of them translated the sentence as (She washed my car). This is because the translators realize that the attached pronoun (تاء) is (تاء التأنيث الساكنة) and is used to specify the identity of the subject as being feminine, neither neutral nor masculine.

The suggested translations for these STs are as follows:

1. She heard him speaking well of you.
2. She washed my car.

1.6. Conclusions:

Arabic language has a unique grammatical structure, which can be understood either ways, with or without diacritics. However, in order to identify the intended syntactic and semantic structures, the sentences must be diacritized. Without vocalization/ diacritization, the sentences can be interpreted as having two or more possible structures, but the correct one is only identified through using diacritics. Through the analysis of data, it is found that translators did not realize the intended syntactic and semantic structures of the sentences when they were presented without diacritics. After diacritizing the sentences, most of them succeeded in proving the appropriate translation for the intended structures. In pre-test, all translators failed to provide the intended structures of sentences under study. Whereas in post-test, not all sentences were translated appropriately, except for ST1 (انتُخِبَ الطالبُ قدوةً), ST2 (غُيِنَ احمدٌ مديراً للشركةِ) and ST6 (غَسَلَتْ سَيارَتي) as all of their translations provided the intended syntactic and semantic structures. As for the other three sentences, namely ST3 (سَجِّلْ أسماءَ الطلبةِ الحاضرين في الدرسِ) ST4 (حَفَظْتَ القصيدةَ؟) and ST5 (سَمِعَتْهُ يتكلمُ عنكَ بخيرٍ), almost all the translators failed in providing the intended structures. Nevertheless, it is clear that vocalization/

diacritization has a great role to play in determining the intended syntactic and semantic structures. Therefore, for avoiding confusion and ambiguity, Arabic sentences must be presented with diacritics in case of translating written texts in order for the translator to know the intended functions and forms of sentence elements, and thus provide appropriated translations for them.

References:

Abu-Rabia, S. (2002). Reading in a root-based-morphology language: the case of Arabic. Journal of Research in Reading, 25,299-309.

Abu-Rabia, S. (2007). The role of morphology and short vowelization in reading Arabic among normal and dyslexic readers in grades 3, 6, 9, and 12. J. Psycholinguist Res., 36 (2), pp. 89-106.

Diab, M., Ghoneim, M. & Habash, N. (2007). Arabic diacritization in the context of statistical machine translation. Proceedings of The Eleventh Machine Translation Summit (MT-Summit XI), Copenhagen, Denmark.

Hallberg, A. (2016). Case endings in Spoken Standard Arabic: Statistics, norms, and diversity in unscripted formal speech. Doctoral dissertation, Lund University.

Hallberg, A. (2022). Principles of variation in the use of diacritics (taškīl) in Arabic books. Language Sciences, 93.

Maroun, M. (2017). Diacritics and the Resolution of Ambiguity in Reading Arabic. PhD dissertation, Department of Psychology, University of Essex, United Kingdom.

Morag, S. (1972). The vocalization systems of Arabic, Hebrew, and Aramaic: Their phonetic and phonemic principles (2nd ed.). The Netherlands: MOUTON & CO • 'S-GRAVENHAGE.

Ryding, K. (2021). Case in Arabic. In K. Ryding, D. Wilmsen (Eds.), The Cambridge Handbook of Arabic Linguistics, Cambridge University Press (2021), pp. 353-370, doi:10.1017/9781108277327.016

Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014). The structure of Arabic language and orthography. In E. Saiegh-Haddad, R.M. Joshi (Eds.), Handbook of Arabic Literacy: Insights and Perspectives, Springer, pp. 3-28.

Schultz, E. (2004). A student grammar of Modern Standard Arabic. Cambridge, England: CUP.