Towards a Grammar of the (Non-)Use of Case Inflection in Standard Arabic

# Abstract

Case inflection (CI) has traditionally been regarded as a core part of Arabic grammar, and nouns and adjectives are in grammars described as being inflected for case by default. In practice, however, CI is only applied in a very limited set of situations in most forms of written and spoken Standard Arabic. This article presents a novel interpretation of CI in Standard Arabic in which it is seen as governed by a two-layered system of rules: the first layer specifies whether CI is to be applied, and the second layer specifies the form of the case marking morpheme, if CI to be applied according to the first layer. In this view, nouns and adjectives are in their default form uninflected for case. The rules of where CI is applied are largely uncodified and differ between various oral and written registers. A typology of registers of Standard Arabic for the purposes of describing these differences is presented. This alternative model of CI in Standard Arabic has advantages over the traditional model in that it simplifies the grammatical description and reflects linguistic conventions.

*Keywords:* Arabic; case; registers; modality; morphology; language ideology; grammar

# Introduction

Standard Arabic (SA) is commonly described as a language with morphologically marked case. In authoritative grammars of Arabic we find detailed descriptions of how different classes of nominals are inflected in different syntactic positions. However, it takes only a brief look at some Arabic texts, or a quick listen to someone speaking SA, to see that the endings that mark case on words are in fact very sparse, with a large span of variation between different registers across the written and spoken modes. In news broadcasts, for example, the news anchor will often report a news story with diligent use of case markers on nouns and adjectives, and then turn to a guest in the studio to ask them a question, still in perfectly good SA, but now with no case markers anywhere in the sentence or with case markers used only sporadically. In books commenting on classical poetry or the Koran, as well as in grammars, the passage that is explained or analyzed is typically printed with full diacritization and thus with all case markers present. The accompanying text explaining the passage is however printed without diacritics and therefore with case being marked only sporadically, as is also the case for the vast majority of written text. Statements in grammars such as “Arabic nouns and adjectives normally inflect for three cases”[[1]](#footnote-1) are therefore misleading.

This article puts forward a novel interpretation of the non-use of case inflection (CI) in SA whereby the use of CI is governed by a two-layered system of grammatical rules: one first layer specifying whether or not CI is to be applied on a particular word, and a second layer of rules specifying the phonological and orthographic form of the case marker, if it is to be applied in accordance with the first layer of rules. The uninflected form is in this view regarded to be the default form of nouns and adjectives. On this default for CI is added only under certain conditions. Grammatical descriptions of Arabic have almost exclusively focused on the second layer of rules, relegating the first to domains outside of grammar, typically as depending on speaker proficiency or orthography. These descriptions are thereby sidestepping a question with immense impact on the morphosyntactic structure of SA. The rules of the first layer, of where case inflection is added, differ between various spoken and written registers, and a formalization of these rules therefore requires a typology of registers in SA. A proposal for such a typology is presented below.

The interpretation of the use and non-use of CI as rule governed has both theoretical and pedagogical advantages over the traditional approach. The theoretical advantage lies in a simplification of the description of SA morphosyntax. CI is commonly viewed as a default, ever-present feature of SA. This can be seen in both Badawi, Carter, and Gully’s and Ryding’s grammars, the two most commonly used English language reference grammars of modern SA.[[2]](#footnote-2) These grammars present examples of written language in Arabic script nt which most case markers are absent, but in the accompanying transcription and glossing, case inflection is added. This is illustrated in (1), taken from the grammar of Badawi *et al.* more or less at random.

1. ***tušriqu*** *l-šamsu ʿalā fatarātin mutabāʾidatin*  
   تشرق الشمس على فتارات متباعدة  
   ‘the sun **comes out** at intermittent periods’ (boldface in original)[[3]](#footnote-3)

In the transcription of this example, the case markers *-u, -in*, and *-in* that are not present in the Arabic sample are added to the three nominals respectively. Word internal vowels are also not represented in the Arabic script, but they are tied to the respective lexeme, are retrieved as part of the process of word identification during reading, and must in some form be realized for the sentence to be read aloud. The case markers, on the other hand, are not part of the lexeme and may or may not be added by the person reading. Furthermore, their appearance in both written text and reading aloud is in most context *prohibited*, as detailed later in this article. In effect, example sentences in grammars are presented in a forms that do not occur in the forms of written SA they are meant to represent. In consistently analyzing samples of written language as if CI is present, these grammars add undue complexity to the description of written SA. The more straight-forward interpretation of sentences such as (1) as lacking CI (which is not there in the text to begin with) leads to a less complex, and therefore preferable theoretical framework of grammatical description.

The pedagogical advantages of a rule based approach to the non-use of CI in SA is that it allows teachers and authors of grammars and textbooks to explicitly instruct students not to use case markers where it would not be appropriate. This relieves students from the difficult task of deducing their own rules of where to apply CI, a task made all the more difficult by the contradictions between the importance ascribed to CI in the pedagogical literature and the lack of CI in authentic language input.

Since none of the Arabic vernaculars have a parallel system of CI, it is acquired by Arabic speakers only through formal education. Indeed, case is at the center of grammar instruction in the Arab world, to the extent that “grammar” for a lot of people is simply equated with the case system.[[4]](#footnote-4) It is furthermore taught in a traditional, highly formulaic manner, and often with convoluted forms of explanation, making the grammatical knowledge one acquires in formal schooling difficult to apply in practice.[[5]](#footnote-5) Comments on the low proficiency in applying the case system, even among educated speakers, are common in the literature.[[6]](#footnote-6) This should not be interpreted as indicating low literacy skills, since as a means of syntactic disambiguation CI is almost completely superfluous. The syntactic roles of constituents in the clause are determined by other means, primarily word order and verb agreement.[[7]](#footnote-7)

SA also has a system of morphologically marked mood which in many respects is parallel to that of case: it has no parallel in the vernaculars, is imperfectly mastered by most speakers, is largely absent in writing, and is syntactically superfluous. In the Arabic linguistic tradition, case and mood are regarded as one and the same grammatical system; *iʿraab*, straddling both nominal and verbal syntax and morphology. Much of what is said on case in the following is applicable also to mood, as they have much the same mechanisms of variation. This article is however focused solely on case.

It is important for our discussion on CI in SA to make a theoretical distinction between, on the one hand, the language as it appears in grammatical descriptions, and, on the other, the language as it appears in observed use. The distinction between *norm* and *codification* developed by the Prague school theoreticians is useful in this regard.[[8]](#footnote-8) The norm of a variety is the observed use, while the codification is a set of explicitly formulated linguistic rules of that variety, as found in grammars, lexica, and school books. The codification is an attempt to describe a norm for purposes of teaching and regulation. All language varieties have a norm by definition, but codification is typically associated only with standard varieties. Indeed, a useful definition of a standard variety is that it is a variety provided with a codification sanctioned and enforced by authority.[[9]](#footnote-9) The terminological pair of *codification* and *norm* as coined by the Prague linguistics is somewhat ambiguous in that the former can be interpreted as part of the latter, that is, the codification is often conceived of as one element in a system of norms. In the following I will therefore use the terms *prescription* and *convention*. I take prescription to mean a linguistic rule that is explicitly formulated in grammar books or other works on language that are ascribed authority within the language community. I take convention to mean a linguistic rule that is deduced from empirically observed behavior.

The privileged position of a standard variety (*vis á vis* its related non-standard varieties) is maintained by *standard language ideology*, characterized by “intolerance of optional variability in language”[[10]](#footnote-10). The standard language ideology of the Arabic language community is extremely conservative as it draws on both nationalist and religious ideologies in its efforts to restrain or revert linguistic variation and change.[[11]](#footnote-11). In these efforts the codification is a central, and case has since the inception of the Arabic grammatical tradition been a focal point codification and standardization efforts.[[12]](#footnote-12) Furthermore, in the diglossic language situation of Arabic there are no native speakers of the standard variety[[13]](#footnote-13) whose intuitions or usage could serve as a guide to current usage. This further serves to cement the role of the codification as only legitimate reference point for linguistic correctness. The codification is taken to be the ontological basis of the language, in what Suleiman has called “the *naïve-realist* orientation of the Arabic grammatical tradition.”[[14]](#footnote-14)

This situation where the codification is not allowed to be adapted based on usage has led to considerable differences between prescription and convention in SA. This is especially problematic for language teaching.[[15]](#footnote-15) It led the United Nations *Arab Human Development Report* of 2003 to call for linguistic research leading to “rationalisation of grammar.”[[16]](#footnote-16) Case is perhaps the area of grammar where this is most sorely needed.

Before turning to the use of CI in registers of SA, there is one aspect of oral performance in SA that need be addressed, namely the omission of word endings preceding a pause (Ar. *waqf*), which often means the omission of CI. This practice is carefully codified regarding recitation of the Koran,[[17]](#footnote-17) as well as for poetry in the traditional meters,[[18]](#footnote-18) but is for other types of text more vague. The Arab grammarians had much to say about the morphological effects of pause[[19]](#footnote-19) (much of which is obsolete in modern SA), but very little about where it is to be applied outside the Koran and poetry. In the modern literature, pause is typically “defined (rather vaguely) as an audible break in delivery.”[[20]](#footnote-20) Thus, whenever a person makes a breath pause, the final short vowel and nunation (word final /n/ following the case marking short vowel on indefinite nouns) of the preceding word is omitted. A person reading a text aloud therefore has some liberty in where to insert pauses and apply the associated prescriptively sanctioned omission of word endings. The end of a sentence is however an unequivocal pause position where word ending omission is prescriptively to be applied. For our present purposes, we will restrict our discussion of pause to sentence final position, while keeping in mind that pausal suffix omission is often more broadly applied.

# Case inflection in registers of Standard Arabic

Most of the variation in the extent to which CI is employed in SA can be accounted for by positing a limited set of registers, each with its own set of conventions of how and where case is marked, unmarked, or where CI is optional. This division of SA into registers was inspired by Walters, who mapped out varieties of Arabic on a two-dimensional space where the vertical axis represents the diglossic continuum, from vernacular to standard, and the vertical axis represents the “degree of planning”. The degree of planning ranges from *written* as the most highly planned, through *recited* and *read aloud* to *spoken*, as the least planned.[[21]](#footnote-21) Six registers are proposed here, corresponding to the top half of Walters’ model: diacritized and undiacritized writing, recitation, uninflected and inflected reading aloud, and extemporaneous speech. They can be categorized hierarchically in the two modes of speech and writing as follows:

* written
  + diacritized (1)
  + undiacritized (2)
* oral
  + reading aloud
    - inflected (3)
    - uninflected (4)
  + recitation (5)
  + extemporaneous speech (6)

These registers are to some extent archetypal embodiments of sets of conventions. As such it is not claimed here that each and every sample of SA fits neatly in one register. It is claimed, however, that the vast majority of samples do. Furthermore, since these registers are taken to embody linguistic conventions, samples of language that do not fit with one of these registers are likely to be perceived as strange or odd.

In the rest of this section, these six registers are discussed in turn in regard to their respective conventions of CI. The discussion relies on existing research on the specific register where available. For some registers research is lacking and descriptions are more tentative. It is hoped that future research will fill these gaps and confirm or modify the descriptions presented here. A number of examples that do not fit neatly in the typology of registers developed here will also be discussed.

## The diacritized and the undiacritized written registers

In the Arabic writing system there are two layers of orthographic representation. The first is the layer of letters and punctuation. Letters represent consonants, long vowels and diphthongs. The second layer consists of optional diacritics representing short vowels (*fatḥa, ḍamma, kasra*), absence of a vowel between consonants (*sukūn*), lengthening of consonants (*šadda*), and nunation (*tanwīn*). In principle, the optionality of the layer of diacritics allows for virtually endless variability in what diacritics are used and how often. In practice, however, texts gravitate towards the two opposite poles of no use of diacritics or full use of diacritics. Undiacritized or diacritized text thus form the two main registers of written SA.[[22]](#footnote-22) Pure examples of these two registers are rare: in undiacritized text there are often some diacritics added and in diacritized text there are often some potential diacritics missing. For the purposes of this discussion, we will nevertheless use this simplified binary categorization of text as diacritized or undiacritized.[[23]](#footnote-23)

The choice between the diacritized and the undiacritized register in writing is largely determined by genre. The diacritized register is used only in a limited set of text types, namely in (a) religious source texts, such as the Koran, *ḥadīṯ*, and the Bible; (b) texts considered to be part of the literary heritage or canon, such as classical and much of modern poetry, as well as modern editions of classical prose texts; and (c) texts aimed at language learners, including children’s literature and schoolbooks. The domains of the undiacritized register are best described negatively: it is used for all writing that is none of the above. This includes everything from news articles and novels to mundane every-day written material (receipts, signs, lists of contents on consumer products, instructions for technical appliances, forms, etc.). SA is primarily a written variety, and written material is overwhelmingly in the undiacritized register. The undiacritized register is thus the bread and butter of SA and quantitatively and functionally the most important of all the six registers.

In the undiacritized register only those forms of CI that are represented by letters are present. CI can be divided into two types depending on how the case marking suffix is written: in *orthographic CI* the suffix is represented by one or more letters and therefore appear in the undiacritized register, and in *diacritic CI* the suffix is represented only by diacritics and therefore do not appear in the undiacritized register. This distinction, while based purely on orthography, is of relevance also for the description of conventions of CI in oral registers of SA, as shown below. Orthographic CI appear only on certain specific combinations of declension, case, definiteness, and the form of the stem. Table 1 lists all eight declensions[[24]](#footnote-24) and the conditions under which they take orthographically marked CI. As can be seen in the table, there are only two declensions, sound masculine plural and the dual, where orthographic CI is consistently present, but these only have a nominative/non-nominative morphological distinction. Three of the eight declensions (triptote, defective, and the five nouns) take orthographic CI only under certain conditions. For example, a word in the triptote declension takes orthographic CI only in the accusative, if the word is indefinite and the stem does not end in the letter *tāʾ marbūṭa* (ـة) or the letter *hamza* (ء) preceded by a short or long vowel *a* (ـاء/ـأ). The final three declensions (sound feminine plural, diptote and indeclinable) never take orthographic CI. Furthermore, an orthographically marked three-way distinction of case is only present in the fairly infrequent situations where a word in the ‘five nouns’ declension is in construct state, and when a noun in the triptote declension ends with the letter *hamza* and has an enclitic possessive pronoun. The latter is due to the fact that the letter *hamza* in word medial position, for example when followed by a clitic, changes form depending on surrounding vowels, including case markers. The rules concerning *hamza*’s shape are somewhat esoteric and there is a large degree of variation in actual practice.[[25]](#footnote-25) According to Buckley, when *hamza* is followed by an enclitic pronoun it may be written without a chair (*e.g.* ضوءه ‘his light’), thus disabling the orthographic CI for these words.[[26]](#footnote-26) Orthographic CI on stems with final *hamza* is thus unstable.

As mentioned previously, it is common for undiacritized text to include some occasional diacritics. Diacritics and the vowels they represent can be either *lexical* or *morphosyntactic*. Lexical diacritics are specific to the lexeme and do not vary with syntactic context. Morphosyntactic diacritics, on the other hand, represent suffixes added to the word stem to mark a morphosyntactic feature, case on nouns and adjectives and mood on verbs, and as such vary with syntactic position. The few diacritics that are added to undiacritized text are normally lexical diacritics, used to differentiate homographs. Diacritics are for example added to differentiate active and passive verbs (*e.g.* *ḍaraba* ضرب ‘hit’ and *ḍuriba* ضُرب ‘was hit’[[27]](#footnote-27)), or verbs in forms I and II (*e.g.* *darasa* درس ‘study’ and *darrasa* درّس ‘teach’). These diacritics have no bearing on CI.

It should be clear from this quick overview that CI in undiacritized text is quite limited. According to one calculation, only around 6% of nouns and adjectives are inflected for case in undiacritized text.[[28]](#footnote-28) The system of CI in the undiacritized register is topologically odd in in that it affects only small subset of nominals in a limited set of morphosyntactic contexts, and is quite different from the more canonical case systems of many European languages, such as Latin, German, or Russian.[[29]](#footnote-29)

Table 1: Orthographically marked case

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| declension | marked case | condition |  | example |  |  |  |
|  |  | *Syntactic* | *Stem* | *Nominative* | *Genitive* | *Accusative* |  |
| *Sound m.pl.* | Nom., Acc/Gen. |  |  | مدرسون | مدرسين | مدرسين | ‘teachers (m.)’ |
| *Dual* | Nom., Acc/Gen. |  |  | مدرسان | مدرسين | مدرسين | ‘two teachers’ |
| *Triptote* |  |  |  |  |  |  |  |
| — condition a. | Acc. | Indefinite | Not \<a>/<āʾ> اء/أ or <ah> ة | مدرس | مدرس | مدرساً | ‘[a] teacher’ |
| — condition b. | Nom., Acc., Gen. | Enclitic pronoun | <ʾ> ء | أصدقاؤه | أصدقائه | أصدقاءه | ‘his friends’ |
| *Defective* | Acc. | Indefinite | Not <ah> ة | قاض | قاض | قاضياً | ‘[a] judge’ |
| *‘The five nouns’* | Nom., Acc., Gen. | Construct state | أبوه | أباه | أبيه |  | ‘his father’ |
| *Sound f.pl.* | *none* |  |  | مدرسات | مدرسات | مدرسات | ‘teachers (f.)’ |
| *Diptote* | *none* |  |  | مدارس | مدارس | مدارس | ‘schools’ |
| *Indeclinable* | *none* |  |  | فتى | فتى | فتى | ‘[a] boy’ |

In the diacritized register both lexical and morphosyntactic diacritics are supplied. This means that in the diacritized register, CI are present on virtually all nouns and adjectives. Here, unlike the undiacritized registers, we have a full-fledged system of morphologically marked case. The diacritized register is thereby not only orthographically different from the undiacritized register, it is also morphologically different in having a significantly more complex nominal morphology.

A unique feature of the diacritized register is the use of diacritic CI in sentence final position. In the undiacritized register no diacritic CI is, and in the oral registers, including recitation and the most careful forms of reading aloud, they are not enunciated in sentences final position, as described later in this section.

The two registers of undiacritized means that there are by and large only two options for the use of CI in writing: only orthographic CI or complete use of CI. As a means of illustration, consider examples (2–5), the headline of an article on the bbc Arabic news site at the time of writing.[[30]](#footnote-30) This sentence appeared on the news site in the undiacritized written register as (2). None of the six nouns and adjectives in the sentence are inflected for case. What is more, the author or editor does not have the option of inflecting words for case since this would entail using the diacritized register (3), or, conceivably, using only morphosyntactic diacritics (4), which for this domain would be highly unconventional.

1. أبرز الأرقام القياسية في تاريخ كأس العالم  
   most.important def-numbers def-comparable in history cup def-world
2. أَبْرَزُ الأَرْقامِ الْقِياسِيّةِ فِي تَارِيخِ كَأْسِ الْعَالَمِ  
   most.important-nom def-numbers-gen def-comparable-gen in history-gen cup-gen def-world
3. أبرزُ الأرقامِ القياسيةِ في تاريخِ كأسِ العالمِ  
   most.important-nom def-numbers-gen def-comparable-gen in history-gen cup-gen def-world-gen
4. أَبْرَز الأَرْقام الْقِياسِيّة فِي تَارِيخ كَأْس الْعَالَم  
   most.important def-numbers def-comparable in history cup def-world  
   ‘The most important records in the history of the World Cup’

To take another example, a formal professional letter or e-mail would normally be written in SA in the undiacritized register. Inflecting words for case would entail using the diacritized register, in analogy with (3). The use of diacritic case markers in this context would not breach prescriptive grammar, but it would breach linguistic convention. In effect, the author does not have the option to inflect words for case.

Intermediate forms where diacritization is only partially applied does occur, but is uncommon. Text with only morphosyntactic diacritics, as in (4), can be found in grammars[[31]](#footnote-31) and, albeit inconsistently, children’s books.[[32]](#footnote-32) With regards to CI, this is of course equivalent to the diacritized register in that diacritic CI is fully applied. The inverse, text with only lexical diacritics, as in (5), is even more rare. It can be found in some entry level Arabic textbooks.[[33]](#footnote-33)

## Inflected and uninflected reading aloud

Reading aloud is the oral reproduction of a written text based on the written visual input. Only reading aloud of undiacritized text will be discussed here. In reading an undiacritized text aloud, short vowels cannot be retrieved from the text via direct visual input but must be retrieved by other means. Lexical and morphosyntactic vowels are retrieved by very different processes. Lexical vowels are accessed through lexical retrieval and is often necessary to make the written word at all pronounceable. To add lexical vowels, the reader identifies the word from the undiacritized written form and produce the complete phonological form of this word as stored in the mental lexicon.[[34]](#footnote-34) Morphosyntactic vowels and nunation, on the other hand, need not be enunciated for a text to be read aloud. The presence or absence of morphosyntactic vowels and nunation is the main distinguishing feature between what is here identified as the two registers of *inflected* and *uninflected* reading aloud. In inflected reading aloud all, or most diacritic case markers (that are not present in the text) are enunciated, and in the uninflected register they are not. As with the diacritized and undiacritized written registers, the assumption here is that performances of reading aloud cluster around enunciating no or very few diacritic case markers on the one hand, or enunciating all or most diacritic case markers on the other, and that most performances of reading aloud can therefore be classified as one of the two. Uninflected reading aloud, where the reader only enunciates the orthographic case markers already present in the text, is the default, go-to form of reading aloud. It is the register used when a text is read aloud in informal situations where the focus is on the informational contents of the text rather than on its linguistic form. This is for example the case when a person is reading a particularly interesting piece of an article to someone, or when two students review some text and read parts of it aloud in order to discuss it. The register of undiacritized reading aloud is also sometimes used in more formal situations, for example in televised debate and discussion programs when participants read parts of documents aloud as part of their argument.[[35]](#footnote-35) Lectures delivered from a prepared manuscript are also often in this register.

There are to the best of my knowledge no linguistic studies of uninflected reading aloud. This is problematic for the purposes of this article in that we have very little direct, systematic empirical evidence of how CI is employed in this register. While one might reasonably assume that orthographic CI is enunciated, we do not have a good picture if it includes some diacritic CI and to what extent. We know from studies of extemporaneously spoken SA that diacritic CI is systematically avoided on definite nouns and in sentence final pause position (see below), and that to the extent that diacritic CI is used, it are used primarily in words with enclitic pronouns. In both extemporaneous speech and inflected reading aloud the speaker/reader has to rely on their own proficiency in the grammatical system to compute and enunciate diacritic CI. We will therefore assume that it plays out in similar ways in these two registers, extrapolating from extemporaneous speech to inflected reading aloud.

CI on words with enclitic pronouns require some further comments. In words with enclitic pronouns the case marker appears between the noun stem and the enclitic pronoun, as in (6).

1. *fī bayt-i-nā/-ka*  
   in house-gen-our/your.ms  
   ‘in our/your house’
2. *fī bayt-nā/-ak*  
   in house-our/your.ms  
   ‘in our/your house’

In this example the noun is in genitive position as a prepositional complement and the case marker *-i* appears between the noun stem and the enclitic pronoun. The corresponding word uninflected for case, as in (7), is often felt to be less acceptable than uninflected words without enclitic pronoun. Indeed, many teachers of Arabic would claim that forms such as (7) are purely vernacular and unacceptable in any form of SA. Given that few people master the case system, the choice in these words is often between an unmarked form or risking incorrect CI. In extemporaneously spoken SA, inflected and uninflected forms in words with enclitic pronoun are used in roughly equal proportions, with individual speakers varying form none to complete and consistent CI.[[36]](#footnote-36) This means that speakers freely use the uninflected forms in words with enclitic pronouns, in contrast to commonly held beliefs. Given our current situation with a lack of research into uninflected reading aloud, we will here extrapolate from our knowledge of extemporaneously spoken SA and make the preliminary assumption that CI in uninflected reading aloud follow the same pattern.

Inflected reading aloud is the oral rendering of a text with consistent use of morphosyntactic vowels in a way that by and large conform with the prescriptive rules of SA. This entails adding case markers that are not present in the text. Inflected reading aloud is strongly associated with formal official situations where there is an expectation of linguistic correctness. This includes news broadcasts, speeches, and voice-overs. Voice acting in animated movies is often in SA, and then invariably in the register of inflected reading aloud with full use of CI.

Reading a text in the register of inflected reading aloud is not the mere oral reproduction of its contents, as is uninflected reading aloud. Rather, it is a way of “performing” the text (in the non-technical meaning of the word), a display of a certain type of linguistic skill that goes above and beyond common reading proficiency. Inflected reading aloud is the prescriptively sanctioned form of reading aloud and is trained and promoted in Arabic classes in the Arab world, as well as in many Western universities. However, due to its limited use outside of Arabic classes, there is little incentive for speakers to develop and maintain proficiency in this register. Even highly educated native speakers of Arabic are typically unable to performs consistently in this register. This has not been systematically studied but is quite obvious to anyone familiar with the Arabic language community.[[37]](#footnote-37) When learning to read, Arabic-speaking children are trained to phonologically realize words in their case marked forms, and only later, and typically without explicit instruction, do they develop reading strategies characteristic of skilled readers in which diacritic word endings are not phonologically encoded and/or enunciated.[*C.f.* @taouka\_cognitive\_2004]

While the register of inflected reading aloud is generally regarded as more correct than the uninflected reading aloud, it is inappropriate in informal situations. We can illustrate this by returning to the headline in (2). If a news anchor were to read this headline in a news broadcast, a highly formal, public situation where prescriptive correctness is expected, it would be read in the register of inflected reading aloud, with CI consistently enunciated except in pause, as in (8). In an informal situation, such as someone seeing this headline on their phone or in a newspaper and calling someones attention to it by reading it aloud, it would be produced in the register of uninflected reading aloud, without diacritic CI, as in (9). The high-flown “performance” register of inflected reading aloud is uncalled-for in this situation and would most likely be interpreted as done in jest.

1. *abraz-u arqām-in qiyāsiyyat-in fi tārīḫ-i kaʾs-i l-ʿālam*  
   most.important-nom numbers-gen comparable-gen in history-gen cup-gen def-world
2. *abraz arqām qiyāsiyya fi tārīḫ kaʾs il-ʿālam*  
   most.important numbers comparable in history cup def-world  
   ‘The most important records in the history of the World Cup.’

An interesting example of inflected and uninflected reading aloud is found in Holes’ analysis of linguistic variation in speeches by the former Egyptian president Ǧamāl ʿAbd al‑Nāṣir (Gamal Abdel Nasser). In one of the analyzed speeches, ʿAbd al‑Nāṣir is speaking to the parliament “in the purest *fuṣħaa*” with all case and mood endings “in what might be described as ‘scriptural style,’”[[38]](#footnote-38) that is the for of Arabic here referred to as inflected reading aloud. On another occasion ʿAbd an‑Nāṣir gives a speech directed to the public where he takes the role of a teacher[[39]](#footnote-39) reading a text on socialism, interspersing it with explanations in Egyptian Arabic. Example (10) is an extract from this speech. Here he does not inflect words for case except for the orthographic case marker *-an* in *ʿamal-an taqaddumiyy-an* ‘progressive action’, in accordance with the description of inflected reading aloud given above. The example is directly followed in the speech by *zayy maa aal il-miṯāq* ‘as the charter says’, a commentary in Egyptian Arabic.

1. *l-ištirākiyya hiya t-targama aṣ-ṣaḥīḥa li-kawn iṯ-ṯawra ʿamal-an taqaddumiyy-an*  
   def-socialism cop def-translation def-correct for-being def-revolution action-acc progressive-acc

‘It is socialism which is the correct interpretation of the revolution in terms of progressive action.’[[40]](#footnote-40)

Holes comments that[[41]](#footnote-41)

since what is being presented is one man’s exegesis rather than the ‘scripture’ itself (as was the case in SPE3 [the other speech]), prosodic features like *iʕraab* [case and mood inflection] which are important to the status of ‘scripture’ if it is to be accepted as such, but which carry no functional load, are unnecessary *and inappropriate*. (Emphasis here.)

Thus, in taking the informal role of a teacher ʿAbd al‑Nāṣir employs another register, uninflected reading aloud, in which diacritic CI is not added to the text. The register of inflected reading aloud, in which he is clearly proficient, would be incompatible with this role.

## Recitation

Recitation is the oral verbatim reproduction of a text from memory. Arabic texts that are memorized are typically those that are ascribed religious or cultural significance, most importantly the Koran, but also *ḥadīṯ*, prayers, and SA poetry and song. The oral register of recitation is closely connected to the diacritized written register in that those texts that are committed to memory are printed with full use of diacritics. In the register of recitation all case markers are produced, except preceding a pause. In terms of CI, recitation is thus very similar to inflected reading aloud. It is nevertheless useful to keep the two apart due to the very different cognitive circumstances under which they are produced. In reading aloud (of undiacritized text) the speaker must constantly be aware of the syntactic position and declension of each word in order to produce the correct case inflected form. In recitation, on the other hand, the speaker does not need to be aware of the grammar, but need simply memorize the correct word forms. It is in other words quite possible to produce a grammatically impeccable recitation while having little or no proficiency in case as a grammatical system. This is indeed common among Arabic speaking Muslim children and by non-Arabic speaking Muslims who learn to recite parts of the Koran.[[42]](#footnote-42) Recitation and inflected reading aloud are thus two distinctly different skills.

Recitation also differs from inflected reading aloud in that it includes archaic forms of CI not normally found in inflected reading aloud, such as the pausal *-ā* on triptote accusatives, nunation on indefinite nouns with a third weak radical preceded by *a* (*e.g.* *hudan* هُدىً ‘guidance’), and CI on proper names.

It was previously noted that texts in the undiacritized register is read aloud in either the inflected or the uninflected register depending on the situation. Texts associated with recitation do not show this variation when orally produced, but are always produced with full CI. Consider (11), the first verse of Imruʾ al-Qays’ (d. 544 ce) *Qifā nabki*, perhaps the most famous verse in all classical Arabic poetry. This verse, as is the case for poetry at large, is read aloud or recited in this form, with case word inflected for case wherever possible. Reading it instead in the uninflected register, as in (12), which would be common for most other types of texts, is unheard-of, and would probably be considered very odd, even comical. The same goes for the Koran and *ḥadīṯ*. These texts are frozen, as it were, in a fully inflected surface form.

1. *qifā nabki min ḏikr-i ḥabīb-in wa-manzil-ī*  
   stop.dua 1pl.weep of memory-gen beloved-gen and-lodging-gen
2. *qifā nabki min ḏikr ḥabīb wa-manzil*  
   stop.dua 1pl.weep of memory beloved and-lodging  
   ‘Halt, friends both! Let us weep, recalling a love and a lodging’[[43]](#footnote-43)

There are two reasons why these texts are treated differently in terms of CI. First, insisting on prescriptively correct renderings is an expression of reverence for the text. Second, the syllabic structure and the associated rhythmic patterns are central aspects of memorization, and variation in CI would alter these patterns. For some words this would also affect the position of lexical the stress within the word. Memorization therefore requires a fixed surface form.

A special kind of recitation is acting in genres that employ SA, most notably drama in historical settings.[[44]](#footnote-44) Acting is a form of recitation in that it is the oral reproduction of a memorized text (the script). As other forms of recitation it features consistent use of diacritic CI, but unlike other forms of recitation it is not associated with text of the literary and religious canon. Acting in SA thus follows conventions of recitation in terms of CI, even though it is made to appear as extemporaneous speech.

## Extemporaneous speech

Extemporaneous speech is unscripted oral language production. The registers discussed so far are based on written text, whether they be the written text itself or an oral reproduction of it. There are, however, situations where a speaker is expected to use SA but where they cannot rely on reproducing a preexisting written text. Instead they speak freely and extemporaneously. Such situations include lectures (where the lecturer does not strictly rely on a manuscript), panel discussions, and news interviews.

With extemporaneously spoken SA there is a problem of definition as to how much non-standard influence we accept while still calling it SA. Some scholars have been of the opinion that speech that lacks case and mood markers are by definition not SA, but rather another variety, typically referred to as Educated Spoken Arabic.[[45]](#footnote-45) This distinction is predicated on the notion that CI is a core, defining feature of SA, and that if it is omitted, the result can therefore not be regarded as SA. With this approach we could not classify text in the undiacritized register, or the reading aloud of these texts in the uninflected register, as SA, since they also feature minimal CI. This is clearly unsatisfactory. In this article, extemporaneously spoken SA is, following Hallberg, defined functionally, as “the most formal register of Arabic extemporaneous speech by proficient, highly educated native speakers of Arabic.”[[46]](#footnote-46) With this definition we expect spoken SA to feature a large measure of variation, seeing to differences between speakers proficiency in different aspects of SA, and potentially different interpretation of what linguistic forms are suitable for a certain level of formality.

The use of CI in extemporaneously spoken SA has been studied by Meiseles,[[47]](#footnote-47) Schulz,[[48]](#footnote-48) Parkinson,[[49]](#footnote-49) and Hallberg,[[50]](#footnote-50) typically as one of a group of features. The material in these studies cover a time span of over fifty years, and it is quite possible that linguistic norms have changed during this time, particularly with the introduction of satellite television in the nineties and the revolution of pan-Arabic news media.[[51]](#footnote-51) The studies by Meiseles and Schulz, conducted before this period, should therefore be interpreted with some caution.

What is clear from these studies is that there is a wide range of variation between individuals in the extent to which they employ CI in their speech. Many speakers never or only very rarely mark words for case. Other speakers use it more, but with rates of CI above around 10% being progressively more rare. Within this variation there are consistent patterns of how the CI that is used is distributed in different morphosyntactic environments, with some positions being favored for CI and others where it is systematically avoided. These patterns make it possible to formulate rules constraining the variation of CI in extemporaneously spoken SA. First, all four studies except Schulz’, in which this parameter is not discussed, identified nouns with enclitic pronouns as being favored for CI. Hallberg calculates the probability of a nominal with an enclitic pronoun as being inflacted for case to 43.9%, compared to the overall average of 7.5%. Some speakers in his study performed near or at sealing level of CI in this position.[[52]](#footnote-52) Second, all studies except Meiseles note a strong preference to mark case where this would be orthographically represented in written text. This is most clearly seen in the indefinite triptote accusatives. Third, similar to the other oral SA registers, diacritic case markers are not used in sentence final pause positions in extemporaneous speech. Fourth, for words with the definite article, diacritic CI is almost completely absent. This has only been shown in Hallberg, the only study to systematically compare CI in different forms of definiteness, but there the patterns is quite striking.[[53]](#footnote-53) It should be noted that while CI is favored in connection with enclitic pronouns and when it is orthographically marked, it is by no means consistent; it only stands out with higher rates of inflection compared to low rates in other positions. In fact, extemporaneous speech is the only SA register where orthographic CI is not used consistently. Clearly, this is because the speaker cannot rely on the visual input of the written text to produce it, but has to rely on their own proficiency in active use CI. Diacritic CI in positions not mentioned here are used very sparingly, if at all, and rarely at rates exceeding one or two percent.

These patterns together encompass the large range of inter-speaker variation in the use of CI in extemporaneous speech, while at the same time setting a ceiling for CI far below the prescriptive ideal. Note also that by these patterns, CI is in any given morphosyntactic position either not used or is optional. In no position in this register is CI applied consistently.

As with the register of uninflected reading aloud, speaking SA with full use of CI in accordance with prescriptive grammar would be to “overdo” it and to flout linguistic conventions. Parkinson gives a detailed description of an extemporaneous performance where the speaker uses consistent, prescriptively correct CI on almost all nouns and adjectives that are not in pause. While this is certainly an impressive demonstration of a rare skill, it is also highly unconventional. As Parkinson notes, “this very performance could only come from someone who is somewhat withdrawn form an engaged role in the more day-to-day society around him.”[[54]](#footnote-54) It is a performance by a person who either does not understand or actively ignores linguistic conventions. Even the interviewer in the study, who had explicitly instructed the interviewee to use SA, had to struggle to hide his annoyance. In this study Parkinson had handpicked speakers from a larger pool of interviews to illustrate the variation in SA performances. Parkinson describes this speaker as “utterly unique”[[55]](#footnote-55) and he should accordingly not be taken to represent a category of speakers.

## Overview

The previous discussion of CI in the various registers of SA is schematically summarized in Table 2. Undiacritized writing, to the far left in the table, is undoubtedly the most important register of SA, both quantitatively, as making up the majority of language produced in SA, and functionally, as the register in which any and all literate persons engage. CI in this register is very limited, being restricted to a few declensions in specific grammatical positions. Diacritized writing is used only on a small set of specialized text types. It is characterized by full use of CI and is the only register where diacritic CI is used in sentence final pause. It is thus grammatically different for the undiacritized register in having a significantly more complex morphology. Text is normally orally produced in the register of uninflected reading aloud. This register can be seen as the oral manifestation of undiacritized writing, preserving the same structure of CI, but with some optional additions of diacritic CI, primarily in connection with enclitic pronouns. In very formal, public situations, undiacritized writing is orally produced in the register of inflected reading aloud, where the reader adds diacritic CI wherever possible except preceding a pause. Few people are ever required to perform in this register. Some texts that are written in the diacritized register are ascribed cultural or religious importance and are commonly committed to memory in order to be recited. All instances of CI that are present in the printed text are enunciated when the text is recited, except in pause. Extemporaneously spoken SA, finally, shows the most variation in CI. This is due to the fact that there is no position in this register where case is consistently inflected; CI is optional in some positions and is not applied elsewhere. To the extent that CI is used in this register it is used primarily where it has an orthographic representation and in connection with enclitic pronouns.

Table 2: Conventions of case inflection in registers of Standard Arabic. Orthographic case inflection are those forms of inflection that in writing are represented with letters (ـاً, ـون/ـين, ـان/ـين) and diacritic case markers those that are represented only by diacritics (ـَ ـً ـُ ـٌ ـِ ـٍ). The middle section of the table shows a more detailed set of conventions for diacritic case inflection in various positions.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | writing |  | recitation | reading aloud |  | extemp. |
|  | *Undiacritized* | *Diacritized* |  | *Inflected* | *Uninflected* |  |
| Orthographic | + | + | + | + | + | o |
| Diacritic | – | + | + | + | r | r |
| —*al-* | – | + | + | + | – | – |
| —encl. pron. | – | + | + | + | o | o |
| —pause | – | + | – | – | – | – |
| Domain | articles | Koran | Koran | news reports | lectures | lectures |
|  | fiction | *ḥadīṯ* | *ḥadīṯ* | voice-overs | informal | panel disc. |
|  |  | technical | Bible | poetry | speeches | interviews |
|  |  | child. lit. | acting | sermons |  |  |
|  |  | poetry |  | voice acting |  |  |

|  |  |
| --- | --- |
| + | = case inflection |
| – | = no case inflection |
| o | = case inflection optional (0–100%) |
| r | = case inflection rare (<2%) |

# Discussion

This article presents a novel interpretation of CI in SA in which there are a set of conventions, or uncodified grammatical rules, governing where case is, and is not, to be marked. Only the positive rules of CI are officially and prescriptively sanctioned and codified. The negative rules of CI are however of no less importance in grammatical descriptions of SA, as the earlier description of registers illustrates. In his classical article on diglossia, Ferguson emphasized the importance of knowing when to use the High and Low varieties, lest one be an “object of ridicule.”[[56]](#footnote-56) The same could be said about the more granular level of registers of SA. This includes knowing when to inflect nominals for case and, equally important, when not to. Using the appropriate register for a given situation and domain, with their respective conventions of CI, is in other words central to communicative competence in SA in both its written and spoken forms.

It is possible that these conventions have arisen from the inability of most users of SA to apply the prescriptive rules of CI. That is, at some point SA with full use of CI may have been the target form in speaking and reading aloud, but few speakers were able to live up to this ideal and therefor had to revert to different patterns of non-use of CI. This gained legitimacy with repeated use and came to be established as a linguistic convention. Full use of CI then seized to be the target form and came to be unsuitable for many domains and situations. Conventions of non-use of CI thus arose as a compromise between prescriptivist ideals and the proficiency of the users of the language. This would explain the conventions whereby recitation maintains full CI. In recitation, the speakers do not themselves apply the rules in online language production, but rather reproduce a memorized string of word forms. Lack of proficiency in the case system is therefore of little hindrance for use of prescriptively correct CI in this register. Whatever the historical background of these conventions, they are now firmly established in their own right and are synchronically an integral part of SA.

The conventions not to inflect words for case is rarely made explicit in the pedagogical literature. This is probably due to the conflicting relation between these conventions and the prescriptive ideal. We may, with a measure of simplification, interpret prescriptive grammar as stating that all nouns and adjectives be inflected for case, and that diacritical case markers should not be enunciated in pause. This prescription aptly captures conventions of CI in diacritized writing, recitation, and inflected reading aloud. These registers are therefore generally considered representing correct language in the language community and are the forms trained and promoted in formal education.

For the undiacritized written register, prescriptive grammar captures conventions of CI only with a major caveat, namely that diacritics be removed. The problem with this caveat is that it in effect abolishes most of the case system. The consequences of this for grammatical description are rarely recognized. On the contrary, the Arabic grammatical tradition and the educational practices that are derived from it go to great lengths in describing the system of CI, but provide little explicit indication of the fact that in most text most of this system is not to be used. Examples that are presented for analysis are universally in the diacritized register, and they are analyzed on the basis that the diacritics are present. The assumption seems to be that case markers are somehow present in text, also when they are not graphically represented, and that the fully diacritized text therefore is the truest representation of the underlying linguistic form. Western grammars of SA, exemplified here by the grammars of Ryding and Badawi *et al.*, follow a similar practice. These grammars mention that diacritic case markers are usually not present in written text,[[57]](#footnote-57) but, as mentioned in the introduction, they analyze examples of text as if case markers are indeed present. This typically implicit assumption of CI as an underlying feature in SA is explicitly formulated by Ryding, who writes that[[58]](#footnote-58)

The Arabic case-ending system consists primarily of short, word-final vowels, *which are invisible* in conventional written Arabic texts. […] The Arabic case system, then, remains mostly hidden from view in written text and is apparent only when the text is read out loud with complete pronunciation of all vowels. (Emphasis in original.)

This way, the undiacritized register is constructed as complying with the prescriptive ideal of complete and consistent CI; all endings are actually there in the text, they are only invisible. (A problem for this traditional theory of case markers in Arabic is of course that CI is typically *not* pronounced when the written text is read aloud and it is not phonologically encoded in silent reading.)

In uninflected reading aloud and extemporaneous speech, conventions of CI are unmistakably in conflict with prescriptive grammar. Both these registers feature negative conventions of CI in highly frequent situations, most notably for diacritic case markers in words with the definite article. Complete use of CI except in pause, as prescriptive grammar dictates, produces speech that is at odds with linguistic conventions. Since these two registers, uninflected reading aloud and extemporaneous speech, do not conform with prescriptive grammar they are generally held to be sub-par SA. This despite them being the socially excepted and indeed the required form of Arabic in many domains and situations. Such contradictions between beliefs about language and linguistic practice are typically dealt with in language communities by a process that Irvine and Gal has labeled *erasure*; the process by which “facts that are inconsistent with the ideological scheme either go unnoticed or get explained away.”[[59]](#footnote-59) The fact that extremely few speakers can actually speak in a way that would be deemed correct is rarely noted, and when it is, it is often explained away as poor linguistic skills by a particular speaker.

The effects of this erasure are most clearly seen in education, where the conventional way of speaking SA and reading aloud are not often not regarded as a legitimate target form. The marginalization of oral proficiency in Arabic education in the Arab world[[60]](#footnote-60) is probably related to the failure to recognize extemporaneously spoken Arabic as its own register with its own grammar that partly deviates from the traditional prescriptive norm. More importantly, the role of the register of uninflected reading aloud in the acquisition of reading skills is often neglected. This register deviates from the prescriptive norm and is therefore subject to erasure. Reading aloud is a first and necessary step in acquiring phonological encoding skills that are then used in silent reading.[[61]](#footnote-61) Skilled reading of Arabic, arguably the end goal when learning to read, is generally assumed not to involve phonological encoding of diacritical marked morphosyntactic endings.[[62]](#footnote-62) The appropriate form for reading aloud as a first step toward efficient silent reading is therefore uninflected reading aloud. With its lack morphosyntactic vowels it represents the phonological from of words encoded in skilled silent reading. It is however often the register of inflected reading aloud that is trained and promoted in entry-level Arabic textbooks at the university level[[63]](#footnote-63) and in first year school books in the Arab world.[[64]](#footnote-64) That is, when students first learn to read they are trained to produce nouns and adjectives in forms inflected for case. This practice risks establishing a cognitively costly habit of phonological computation of morphosyntactic endings that must later be inhibited for effective silent reading, as well as for developing proficiency in the register of uninflected reading aloud. Inflected reading aloud with full use of CI, the register in which many beginning readers are trained, is a specialist skill which most users of the language will never need. It is of little relevance for general reading proficiency.

A similar form of erasure of the register of uninflected reading aloud can be found in experimental studies of reading development in Arabic. The register of uninflected reading aloud has in these studies often not been regarded as its own register, but rather as a sub-standard Arabic resulting from a failure to perform in the prescriptively correct inflected reading aloud. Failure to produce diacritic case markers has accordingly been considered a reading error indicative of lacking reading proficiency.[[65]](#footnote-65) In a review of research on reading Arabic orthography, Abu-Rabia writes that[[66]](#footnote-66)

reading accuracy in Arabic requires vowelizing word endings according to their grammatical function in the sentence, which requires advanced phonological and syntactical ability.

The ability to “vowelize” word ending has in many studies been found not to correlate with reading comprehension, and Abu-Arabia therefore advises that reading aloud tests meant to asses reading skills should focus on reading accuracy (that is, including CI) rather than reading comprehension.[[67]](#footnote-67) Reading proficiency has in other words been measured in relation to the specialist “performance” register of inflected reading aloud, which even few skilled readers master, and not in relation to the functionally more relevant register of uninflected reading aloud.[[68]](#footnote-68)

# Conclusion

This article has presented a critique of traditional accounts of the case system in SA. Traditional accounts present case inflected forms of nominals as the default by systematically presenting the inflected forms as the basis for analysis. This misrepresents the Arabic case system both quantitatively, in that case marked forms are in the most common registers the exception rather than the rule, and qualitatively, in that proficient users of the language for the most part do not process sentences for morphologically marked case, whether in language production or comprehension. For the functionally most important registers of SA, CI is not present in the text, nor in the minds of readers or speakers.

The alternative view of the SA case system presented in this article is that CI is activated only under certain conditions, and is otherwise absent from the language. A descriptive grammar of SA therefore requires not only a description of how different words inflect for case, but also where this inflection is applied in the first place. What these conditions are differ between different spoken and written registers. A proposal for a topology of registers of SA for the purposes of describing these differences was presented in this article. It is hoped that future empirical research will help refine this description.

One effect of regarding nominals without CI the default form of these words is a rearrangement of the status of registers of SA. Traditionally, since CI is regarded as a core feature of SA, and nominals as inflected for case by default, the registers that feature CI most consistently (diacritized writing, inflected reading aloud, and recitation) are taken to represent the canonical form of SA. These registers therefore typically constitute the target form in language instruction, form the basis for grammatical description, and are taken as the reference point for evaluating language skills. If we on the other hand consider SA nominals to be uninflected for case by default, as is suggested in this article, this hierarchy of canonicity is inverted, making registers featuring limited CI (undiacritized writing and uninflected reading aloud) the primary and canonical forms of SA. These registers are in this alternative view the suitable target forms for language instruction, the appropriate basis for grammatical description, and are the relevant reference point for evaluating language skills. This does not imply that the case system should be ignored in teaching and grammar writing, it is indeed a part of certain forms SA. It does imply however that it be treated as a subsidiary grammatical feature, and that the limitations of its application be recognized, formalized and explicated.

1. Karin C. Ryding, *A Reference Grammar of Modern Standard Arabic*, Cambridge, UK: Cambridge University Press, 2005, p. 54. [↑](#footnote-ref-1)
2. El-Said M. Badawi, Michael G. Carter, and Adrian Gully, *Modern Written Arabic: A Comprehensive Grammar*, London: Routledge, 2004; Ryding, *Reference Grammar*. [↑](#footnote-ref-2)
3. Badawi, Carter, and Gully, *Modern Written Arabic*, p. 395. [↑](#footnote-ref-3)
4. Muhammad H. Ibrahim, “Linguistic Distance and Literacy in Arabic,” *Journal of Pragmatics* 7/5 (1983), p. 507–15; Niloofar Haeri, *Sacred Language, Ordinary People: Dilemmas of Culture and Politics in Egypt*, New York: Palgrave Macmillan, 2003, p. 40; Allon J. Uhlmann, “Arabs and Arabic Grammar Instruction in Israeli Universities: Alterity, Alienation and Dislocation,” *Middle East Critique* 21/1 (2012), p. 101–16. [↑](#footnote-ref-4)
5. Ramzi Baalbaki, “Teaching Arabic at University Level: Problems of Grammatical Tradition,” *Proceedings of the Colloquium on Arabic Linguistics, Bucharest, August 29 ê Sept. 2*, 1994, p. 96; Munther Younes, “On ‘I‘rāb’, Power, and Language Reform in the Arab World,” *Al-ʿArabiyya* 40/41 (2007), p. 232ff; Uhlmann, “Arabs and Arabic grammar”; Andreas Hallberg, *Case Endings in Spoken Standard Arabic: Statistics, Norms, and Diversity in Unscripted Formal Speech*, Lund: Lund University, 2016, p. 67ff. [↑](#footnote-ref-5)
6. A. F. L Beeston, *The Arabic Language Today*, London: Hutchinson, 1970, p. 43; Anwar G. Chejne, *The Arabic Language: Its Role in History*, Minneapolis, MN: University of Minnesota Press, 1969, p. 50; Alan S. Kaye, “Remarks on Diglossia in Arabic: Well-Defined vs. Ill-Defined,” *Linguistics* 10/81 (1972), p. 32–48; Dilworth B. Parkinson, “Knowing Standard Arabic: Testing Egyptians’ MSA Abilities,” ed. Mushira Eid and Clive Holes, *Perspectives on Arabic Linguistics*, 5 (1993), p. 47–73. [↑](#footnote-ref-6)
7. Federico C. Corriente, “On the Functional Yield of Some Synthetic Devices in Arabic and Semitic Morphology,” *The Jewish Quarterly Review* 62/1 (1971), p. 20–50; Federico Corriente, “Again on the Functional Yield of Some Synthetic Devices in Arabic and Semitic Morphology (a Reply to J. Blau),” *The Jewish Quarterly Review* 64/2 (1973), p. 154–63; Clive Holes, *Modern Arabic: Structures, Functions, and Varieties*, Washington, DC: Georgetown University Press, 2004, p. 173. [↑](#footnote-ref-7)
8. Bohuslav Havránek, “Zum Problem der Norm in der heutigen Sprachwissenschaft und Sprachkultur,” in *Soziolinguistik: Ansätze zur soziolinguistischen Theoriebildung*, ed. Hugo Steger, Darmstadt: Wissenschaftliche Buchgesellschaft, 1982, p. 289–96; Alois Jedlička, “Theorie der Literatursprache,” in *Grundlagen der Sprachkultur: Beiträge der Prager Linguistik zur Sprachtheorie und Sprachpflege*, ed. Jürgen Scharnhorst and Erika Ising, Berlin: Akademie-Verlag, 1982, p. 40–91. [↑](#footnote-ref-8)
9. Ulrich Ammon, “Explikation der Begriffe ‘Standardvarietät’ und ‘Standardsprache’ auf normtheoretischer Grundlage,” in *Sprachlicher Substandard*, ed. Günter Holtus and Edgar Radke, Tübingen: Niemeyer, 1986, p. 1–63. [↑](#footnote-ref-9)
10. James Milroy and Lesley Milroy, *Authority in Language: Investigating Language Prescription and Standardisation*, London: Routledge, 1991, p. 26. [↑](#footnote-ref-10)
11. Enam Al-Wer, “Arabic Between Reality and Ideology,” *International Journal of Applied Linguistics* 7/2 (1997), p. 251–65; Yasir Suleiman, *The Arabic Language and National Identity: A Study in Ideology*, Washington, DC: Georgetown University Press, 2003; Yasir Suleiman, “The Simplification of Arabic Grammar and the Problematic Nature of the Sources,” *Journal of Semitic Studies* 41/1 (1996), p. 99–119. [↑](#footnote-ref-11)
12. Georges Bohas, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli, *The Arabic Linguistic Tradition*, London: Routledge, 1990; Chejne, *The Arabic Language*; Kees Versteegh, “Arabic Grammar and Corruption of Speech,” *Al-Abhath* 31 (1983), p. 139–60. [↑](#footnote-ref-12)
13. Charles A. Ferguson, “Diglossia,” *Word* 15 (1959), p. 325–40; Charles A. Ferguson, “Epilogue: Diglossia Revisited,” in *Understanding Arabic: Essays in Contemporary Arabic Linguistics in Honor of El-Said Badawi*, ed. Alaa Elgibali, Cairo: American University in Cairo Press, 1996, p. 49–67. [↑](#footnote-ref-13)
14. Suleiman, “The simplification of Arabic,” p. 114. [↑](#footnote-ref-14)
15. Parkinson, “Knowing Standard Arabic”; Mohamed Maamouri, “Language Education and Human Development: Arabic Diglossia and Its Impact on the Quality of Education in the Arab Region.” Philadelphia: International Literacy Institute, 1998; Kassem M. Wahba, “First Language Teaching,” in *Encyclopedia of Arabic Language and Linguistics*, ed. C. H. M. Versteegh, vol. 2 Leiden: Brill, 2006. [↑](#footnote-ref-15)
16. F. Bennani et al., “Arab Human Development Report 2003: Building a Knowledge Society” New York: United Nations Development Program, 2003, p. 125. [↑](#footnote-ref-16)
17. Kristina Nelson, *The Art of Reciting the Qurʾan*, Cairo; New York: The American University in Cairo Press, 2001, p. 19, 28–29. [↑](#footnote-ref-17)
18. Harris Birkeland, *Altarabische Pausalformen*, Oslo: I kommisjon hos J. Dybwad, 1940. [↑](#footnote-ref-18)
19. For details, see William Wright, *A Grammar of the Arabic Language*, New York: Cosimo, 2011, 358, vol.2 [↑](#footnote-ref-19)
20. Holes, *Modern Arabic*, p. 63. [↑](#footnote-ref-20)
21. Keith Walters, “Fergie’s Prescience: The Changing Nature of Diglossia in Tunisia,” *International Journal of the Sociology of Language* 163/1 (2003), p. 97. [↑](#footnote-ref-21)
22. These registers are often referred to as “vowelled” and “unvowelled” text. These designations are problematic, first, in that the diacritics represent not only vowels, and second, in that they conflate phonetic entities (vowels) with graphic entities (diacritics). [↑](#footnote-ref-22)
23. Some scholars have documented a variety of the undiacritized register, Informal Written Arabic, which not included in the typology of registers developed here. Informal Written Arabic is text that is SA at its base but in which adherence to certain specifically SA features, such as CI, is relaxed. Gustav Meiseles, “Informal Written Arabic: A Preliminary Evaluation of Data,” *Israel Oriental Studies* 6 (1979), p. 73–313; R. Kirk Belnap and Brian Bishop, “Arabic Personal Correspondence: A Window on Change in Progress?” *International Journal of the Sociology of Language* 163/1 (2003), p. 9–25. This register is described as being primarily used in personal written correspondence, the social function of which is today increasingly filled by electronic communication in social media and instant text messaging written in the dialects. Dialectal writing is gradually gaining legitimacy in a number of domains and its orthography is increasingly conventionalized. David Palfreyman and Muhamed al Khalil, “‘A Funky Language for Teenzz to Use’: Representing Gulf Arabic in Instant Messaging,” *Journal of Computer-Mediated Communication* 9/1 (2003); Madiha Doss and Humphrey Davies, eds., *al-ʿĀmmiyya l-miṣriyya l-maktūba*, Cairo: al-Hayʾa l-miṣriyya l-ʿāmma li-l-kitaab, 2013; Kristian Takvam Kindt, Jacob Høigilt, and Tewodros Aragie Kebede, “Writing Change: Diglossia and Popular Writing Practices in Egypt,” *Arabica* 63/3 (2016), p. 324–76; Jacob Høigilt and Gunvor Mejdell, eds., *The Politics of Written Language in the Arab World*, Leiden; Boston: Brill, 2017. There is thus an emerging binary division of labor where informal writing is in the dialects and thus unbound by prescriptive norms of SA, while formal writing is in SA with prescriptively correct language as the target form. Informal writing in the dialect is not a register of SA, and thus fall outside of the scope of this article. In this binary division, Informal Written Arabic, in the sense of SA where prescriptively correct language is not the target form, is marginalized. [↑](#footnote-ref-23)
24. Ryding, *Reference Grammar*, p. 184ff. [↑](#footnote-ref-24)
25. Timothy Buckwalter, “Issues in Arabic Morphological Analysis,” in *Arabic Computational Morphology*, ed. Abdelhadi Soudi, Antal van den Bosch, and Günter Neumann, 38 Dordrecht: Springer, 2007, p. 23–41; Dilworth B. Parkinson, “Orthographic Variation in Modern Standard Arabic: The Case of the Hamza,” ed. Mushira Eid and John McCarthy, *Perspectives on Arabic Linguistics*, 2 (1990), p. 269–95. [↑](#footnote-ref-25)
26. Ronald P. Buckley, *Modern Literary Arabic: A Reference Grammar*, 1st ed. Beirut: Librairie du Liban Publishers, 2004, p. 15. [↑](#footnote-ref-26)
27. *Cf.* Ehab W. Hermena et al., “Processing of Arabic Diacritical Marks: Phonological/Syntactic Disambiguation of Homographic Verbs and Visual Crowding Effects,” *Journal of Experimental Psychology: Human Perception and Performance* 41/2 (2015), p. 494–507. [↑](#footnote-ref-27)
28. Hallberg, *Case Endings*, p. 76n40. [↑](#footnote-ref-28)
29. Laura Janda, “Inflectional Morphology,” in *The Oxford Handbook of Cognitive Linguistics*, ed. Dirk Geeraerts and Hubert Cuyckens, 2010, p. 632–49; Greville G. Corbett, “Determining Morphosyntac Features Values: The Case of Case,” in *Case and Grammatical Relations: Studies in Honor of Bernard Comrie*, ed. Greville G. Corbett and Michael P. Noonan, John Benjamins Publishing, 2008, p. 1–43. [↑](#footnote-ref-29)
30. <http://www.bbc.com/arabic/sports-44130063>. Accessed May 25, 2018. [↑](#footnote-ref-30)
31. Muhammad al-Sharkawi, “The Ecology of Case in Modern Standard Arabic,” *Folia Orientalia* LIII (2016), p. 249. [↑](#footnote-ref-31)
32. *E.g.* Astrid Lindgren, *Īmīl fī Lunnabaryā yasraḥ wa-yamraḥ*, Djursholm: Dar Al-Muna, 2008; Taġrīd al-Naǧǧār and Ǧulnār Ḥājū, *Sitt al-kull*, Amman, Jordan: Salwa Publishers, 2015. [↑](#footnote-ref-32)
33. *E.g.* al-Saʿīd M. Badawī and Fatḥī ʿAlī Yūnis, *al-Kitāb al-asāsī fī taʿlīm al-luġa l-ʿarabiyya li-ġayr al-nāṭiqīn bihā*, Tunis: al-Munaẓẓama l-ʿrabiyya li-l-tarbiya wa-l-ṯaqāfa wa-l-ʿulūm, 1973. [↑](#footnote-ref-33)
34. Note that the lexical vowels supplied when reading a word aloud need not necessarily be the vowels prescribed in SA. The proper name *muḥammad* (محمد <mḥmd>) could for example be stored in the mental lexicon and read aloud with a dialectal vowel pattern as *mḥammed*. [↑](#footnote-ref-34)
35. See for instance around the three and a half minute mark in the episode of AlJazeera’s program *al-Ittiǧāh al-muʿākis* broadcast 26 December 2017, available on AlJazeera’s YouTube channel (<https://www.youtube.com/watch?v=U9Yec92LOyo>, accessed 14 November 2019). Here the speaker goes from speaking extemporaneously with SA syntax and lexicon but no CI, to reading aloud from a piece of paper, enunciating only case markers that are orthographically represented in the text, as well diacritic case markers on nouns with enclitic pronouns, some prescriptively incorrect. [↑](#footnote-ref-35)
36. Hallberg, *Case Endings*, p. 205. [↑](#footnote-ref-36)
37. Beeston, *The Arabic Language Today*, p. 53; Ibrahim, “Linguistic distance”; Kaye, “Remarks on diglossia in Arabic”; Elinor Saiegh-Haddad and Rachel Schiff, “The Impact of Diglossia on Voweled and Unvoweled Word Reading in Arabic: A Developmental Study from Childhood to Adolescence,” *Scientific Studies of Reading* 20/4 (2016), p. 1–14. [↑](#footnote-ref-37)
38. Clive Holes, “The Uses of Variation: A Study of the Political Speeches of Gamal Abd Al-Nasir,” ed. Mushira Eid and Clive Holes, 5 (1993), p. 29. [↑](#footnote-ref-38)
39. *Ibid.*, p. 29. [↑](#footnote-ref-39)
40. *Ibid.*, p. 29. [↑](#footnote-ref-40)
41. *Ibid.*, p. 31. [↑](#footnote-ref-41)
42. Haeri, *Sacred Language, Ordinary People*, p. 41; Daniel A. Wagner and Abdelhamid Lotfi, “Learning to Read by ‘Rote’,” *International Journal of the Sociology of Language* 1983/42 (1983), p. 111–21; Kees Versteegh, “Learning Arabic in the Islamic World,” in *The Foundations of Arabic Linguistics III: The Development of a Tradition: Continuity and Change*, ed. Georgine Ayoub and Kees Versteegh, Boston, MA: Brill, 2018, p. 245–67. [↑](#footnote-ref-42)
43. Translation from A. J. Arberry, *The Seven Odes: The First Chapter in Arabic Literature*, London: Routledge, 2018, p. 61. [↑](#footnote-ref-43)
44. Clive Holes, “Orality, Culture, and Language,” in *The Oxford Handbook of Arabic Linguistics*, ed. Jonathan Owens, Oxford: Oxford University Press, 2013, p. 281–99. [↑](#footnote-ref-44)
45. Terence Frederick Mitchell, “What Is Educated Spoken Arabic?” *International Journal of the Sociology of Language* 61 (1986), p. 19. [↑](#footnote-ref-45)
46. Hallberg, *Case Endings*, p. 6. [↑](#footnote-ref-46)
47. Gustav Meiseles, “Restitution of ‘Word-Endings’ in Modern Literary Arabic,” *Israel Oriental Studies* 7 (1977), p. 173–95. [↑](#footnote-ref-47)
48. David Eugene Schulz, *Diglossia and Variation in Formal Spoken Arabic in Egypt*, Ann Arbor: Unpublished Ph.D. thesis, 1981. [↑](#footnote-ref-48)
49. Dilworth B. Parkinson, “Speaking Fuṣḥā,” in *Arabic Sociolinguistics: Issues and Perspectives*, ed. Yasir Suleiman, Surrey: Curzon Press, 1994, p. 179–211. [↑](#footnote-ref-49)
50. Hallberg, *Case Endings*. [↑](#footnote-ref-50)
51. Andrew Hammond, *Popular Culture in the Arab World: Arts, Politics, and the Media*, Cairo: American University in Cairo Press, 2007, p. 209ff. [↑](#footnote-ref-51)
52. Hallberg, *Case Endings*, p. 208f. [↑](#footnote-ref-52)
53. *Ibid.*, p. 209. [↑](#footnote-ref-53)
54. Parkinson, “Speaking fuṣḥā,” p. 209. [↑](#footnote-ref-54)
55. *Ibid.*, p. 207. [↑](#footnote-ref-55)
56. Ferguson, “Diglossia,” p. 239. [↑](#footnote-ref-56)
57. Ryding, *Reference Grammar*, p. 166; Badawi, Carter, and Gully, *Modern Written Arabic*, p. 20f. [↑](#footnote-ref-57)
58. Ryding, *Reference Grammar*, p. 166. [↑](#footnote-ref-58)
59. Judith T. Irvine and Susan Gal, “Language Ideology and Linguistic Differentiation,” in *Linguistic Anthropology: A Reader*, ed. Alessandro Duranti, 1 Malden, Mass: Blackwell Publishers, 2001, p. 404. [↑](#footnote-ref-59)
60. Wahba, “First language teaching.” [↑](#footnote-ref-60)
61. Alexander Pollatsek, “The Role of Sound in Silent Reading,” in *The Oxford Handbook of Reading*, ed. Alexander Pollatsek and Rebecca Treiman, New York: Oxford University Press, 2015, p. 185–201; Keith Rayner et al., *Psychology of Reading*, New York; London: Psychology Press, 2012, chap. p. 7. [↑](#footnote-ref-61)
62. Corriente, “On the functional yield of some synthetic devices in arabic and semitic morphology”; Jaroslav Stetkevych, *The Modern Arabic Literary Language: Lexical and Stylistic Developments*, Washington DC: Georgetown University Press, 2006; Mary Catherine Bateson, *Arabic Language Handbook*, Washington, D.C: Center for Applied Linguistics, 1967; Saiegh-Haddad and Schiff, “Impact”; Miriam Taouka and Max Coltheart, “The Cognitive Processes Involved in Learning to Read in Arabic,” *Reading and Writing* 17/1 (2004), p. 27–57. [↑](#footnote-ref-62)
63. *E.g.* Bassam K. Frangieh, *Arabic for Life: A Textbook for Beginning Arabic*, New Haven, Conn.: Yale University Press, 2012; Eckehard Schulz, Günther Krahl, and Wolfgang Reuschel, *Standard Arabic: An Elementary/Intermediate Course*, Cambridge: Cambridge University Press, 2000; Mahdi Alosh, *Ahlan Wa-Sahlan: Functional Modern Standard Arabic for Beginners*, Yale Language Series New Haven, Conn: Yale University Press, 2000. [↑](#footnote-ref-63)
64. *E.g.* Ḥulmī Ḥasan et al., *Luġatunā l-jamīla*, ed. ʿUmar Muslim, Ramallah, Palestine: Wizārat al-tarbiya wa-l-taʿlīm al-ʿālī, 2004; Muḥammad Ṣalāḥ Farj and Muḥammad ʿAbd al-Ḥamīd Ġurāb, *Hayyā natawāṣal*, Cairo: Wizārat al-tarbiya wa-l-taʿlīm, qiṭāʿ al-kutub, 2011; Firās Ḥūš and ʿAzzat Talja, eds., *al-ʿArabiyya luġatī*, Damascus: Wizārat at-tarbiya, 2011. [↑](#footnote-ref-64)
65. Salim Abu-Rabia, “Reading Arabic Texts: Effects of Text Type, Reader Type and Vowelization,” *Reading and Writing* 10/2 (1998), p. 105–19; Salim Abu-Rabia, “The Role of Vowels in Reading Semitic Scripts: Data from Arabic and Hebrew,” *Reading and Writing* 14/1 (2001), p. 39–59; Raphiq Ibrahim, Zohar Eviatar, and Judith Aharon-Peretz, “Metalinguistic Awareness and Reading Performance: A Cross Language Comparison” 36/4 (2007), p. 297–317. [↑](#footnote-ref-65)
66. Salim Abu-Rabia and Haitham Taha, “Reading in Arabic Orthography: Characteristics, Research Findings, and Assesment,” in *Handbook of Orthography and Literacy*, ed. R. Malatesha Joshi and P. G. Aaron, Mahwah, N.J.: L. Erlbaum Associates, 2006, p. 321. [↑](#footnote-ref-66)
67. *Ibid.*, p. 336. [↑](#footnote-ref-67)
68. Saiegh-Haddad and Schiff, “Impact.” and Taouka and Coltheart, “Cognitive processes.” are examples of experimental studies of reading proficiency where production of case markers did not affect proficiency scores. [↑](#footnote-ref-68)